

58447

From: Chan, Christina
Sent: Wednesday, January 16, 2002 1:54 PM
To: Li, Janice; STIC-Biotech/ChemLib
Subject: RE: rush search for 09/242,202

Please rush. Thanks Chris

-----Original Message-----

Fr m: Li, Janice
Sent: Wednesday, January 16, 2002 1:26 PM
T : Chan, Christina
Subject: rush search for 09/242,202

Hi, Chris;

Could you please give approval for rush search SEQ ID Nos: 10, 16, 22, 27, 28; residues 1-12 of SEQ ID No: 22; and residues 1-11 of SEQ ID No: 10 against all commercial databases including pending patent database.

This is an amended case.

Thanks,

Q. Janice Li
Patent Examiner
AU 1632
CM1, Rm12D11
Mail Box12E12
703-308-7942
janice.li@uspto.gov

Point of Contact:
Mona Smith
Technical Info. Specialist
CM1 12C14 Tel: 308-3278

RECEIVED
JAN 16 2002
STIC

Searcher: M. Smith
Phone: _____
Location: _____
Date Picked Up: 1/17/02
Date Completed: 1/18
Searcher Prep/Review: 15
Clerical: _____
Online time: 15

TYPE OF SEARCH:
NA Sequences: 7
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

10/11/1994
10/11/1994
10/11/1994
10/11/1994

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:54:39 ; Search time 221.34 Seconds
(without alignments)
12.279 Million cell updates/sec

Title: US-09-242-202A-22_COPY_1_12
Perfect score: 12
Sequence: 1 GCCACCATGGCC.12

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/PTOUS_COMB.seq.*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	100.0	15	US-09-309-382-9	Sequence 9, Appli
2	12	100.0	18	US-09-309-382-8	Sequence 8, Appli
3	12	100.0	22	US-09-309-382-24	Sequence 24, Appl
4	12	100.0	22	US-09-309-382-25	Sequence 25, Appl
5	12	100.0	27	US-08-013-801-8	Sequence 8, Appli
6	12	100.0	27	US-08-072-063-17	Sequence 17, Appl
7	12	100.0	27	US-08-212-132-8	Sequence 8, Appli
8	12	100.0	27	US-08-064-693-17	Sequence 17, Appl
9	12	100.0	27	US-08-430-417-8	Sequence 8, Appli
10	12	100.0	27	US-08-470-366-8	Sequence 8, Appli
11	12	100.0	27	US-08-466-822-8	Sequence 8, Appli
12	12	100.0	27	US-08-704-504-8	Sequence 8, Appli
13	12	100.0	27	US-08-885-366-17	Sequence 17, Appl
14	12	100.0	27	US-09-223-342-8	Sequence 8, Appli
15	12	100.0	27	PCT-US93-04754-17	Sequence 17, Appl
16	12	100.0	27	PCT-US94-01235-8	Sequence 8, Appli
17	12	100.0	27	PCT-US95-03125-8	Sequence 8, Appli
18	12	100.0	28	US-08-227-536-13	Sequence 13, Appl
19	12	100.0	28	US-09-309-382-17	Sequence 17, Appl
20	12	100.0	28	US-09-309-382-18	Sequence 18, Appl
21	12	100.0	28	PCT-US95-04682-13	Sequence 13, Appl
22	12	100.0	32	US-08-987-418A-9	Sequence 9, Appli
23	12	100.0	32	US-09-343-062-9	Sequence 9, Appli
24	12	100.0	33	US-08-478-386A-4	Sequence 4, Appli
25	12	100.0	33	US-08-292-597-4	Sequence 4, Appli
26	12	100.0	33	US-08-388-653-4	Sequence 4, Appli
27	12	100.0	33	US-08-473-985-4	Sequence 4, Appli

28	12	100.0	33	US-08-483-898-4	Sequence 4, Appli
29	12	100.0	33	US-09-087-716-4	Sequence 4, Appli
30	12	100.0	33	US-09-157-753-4	Sequence 4, Appli
31	12	100.0	33	US-09-157-230-4	Sequence 4, Appli
32	12	100.0	33	US-09-087-811-4	Sequence 4, Appli
33	12	100.0	33	US-09-156-855-4	Sequence 4, Appli
34	12	100.0	33	US-09-158-010-4	Sequence 4, Appli
35	12	100.0	33	US-09-087-647-4	Sequence 4, Appli
36	12	100.0	34	US-08-264-003B-6	Sequence 6, Appli
37	12	100.0	34	US-08-842-234-6	Sequence 6, Appli
38	12	100.0	35	US-09-231-182B-3	Sequence 3, Appli
39	12	100.0	37	US-09-309-382-20	Sequence 20, Appl
40	12	100.0	38	US-09-231-182B-4	Sequence 4, Appli
41	12	100.0	39	US-09-309-382-21	Sequence 21, Appl
42	12	100.0	39	US-09-309-382-22	Sequence 22, Appl
43	12	100.0	40	US-08-425-684-29	Sequence 29, Appl
44	12	100.0	40	US-08-425-684-32	Sequence 32, Appl
45	12	100.0	40	US-08-675-502-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-309-382-9
; Sequence 9, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: n bases may be A, T, C, G, unknown or other
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide
US-09-309-382-9

Query Match 100.0%; Score 12; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
Db 4 gccaccatggcc 15

RESULT 2
US-09-309-382-8
; Sequence 8, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recognition
; OTHER INFORMATION: site
US-09-309-382-8

Query Match 100.0%; Score 12; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 2 gccaccatggcc 13

RESULT 3
US-09-309-382-24
; Sequence 24, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: n bases may be A, T, C, G, unknown or other
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide
US-09-309-382-24

Query Match 100.0%; Score 12; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 9 gccaccatggcc 20

RESULT 4
US-09-309-382-25/c
; Sequence 25, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 22

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: n bases may be A, T, C, G, unknown or other
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide
US-09-309-382-25

Query Match 100.0%; Score 12; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 14 GCCACCATGGCC 3

RESULT 5
US-08-013-801-8
; Sequence 8, Application US/08013801
; Patent No. 5420019
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Horwitz, Arnold
; APPLICANT: Burke, David
; APPLICANT: Baltayan, Manik
; APPLICANT: Grinna, Lynn S
; TITLE OF INVENTION: Stable Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Products and Pharmaceutical
; TITLE OF INVENTION: Compositions Containing the Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,801
; FILING DATE: 02 FEB 1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/346-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-013-801-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12

COUNTRY: United States of America

SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-064-693-17

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACCATGGCC 12
Db 10 GCCACCATGGCC 21

RESULT 9
US-08-430-417-8
; Sequence 8, Application US/08430417
; Patent No. 5674834
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Horwitz, Arnold
; APPLICANT: Burke, David
; APPLICANT: Baltalan, Manik
; APPLICANT: Grinna, Lynn S.
; TITLE OF INVENTION: Stable Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein Products and Pharmaceutical Compositions Containing
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430,417
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-430-417-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACCATGGCC 12
Db 10 GCCACCATGGCC 21

RESULT 10
US-08-470-366-8
; Sequence 8, Application US/08470366
; Patent No. 5703038
; GENERAL INFORMATION:
; APPLICANT: Little, Roger
; APPLICANT: Ammons, Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Dimer Products
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,366
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-470-366-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
Db 10 GCCACCATGGCC 21

RESULT 11
US-08-466-822-8
; Sequence 8, Application US/08466822
; Patent No. 5827816
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Horwitz, Arnold
; APPLICANT: Burke, David
; APPLICANT: Baltalan, Manik
; APPLICANT: Grinna, Lynn S.
; TITLE OF INVENTION: Stable Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein Products and Pharmaceutical Compositions Containing
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago

; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,822
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-466-822-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
DB 10 GCCACCATGGCC 21

RESULT 12
US-08-704-504-8
; Sequence 8, Application US/08704504
; Patent No. 5856302
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; APPLICANT: Ammons, William Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Dimer Products
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,504
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,132
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-704-504-8

Query Match 100.0%; Score 12; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
DB 10 GCCACCATGGCC 21

RESULT 13
US-08-885-366-17
; Sequence 17, Application US/08885366
; Patent No. 6274348
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Grinna, Lynn S
; APPLICANT: Horwitz, Arnold
; TITLE OF INVENTION: BPI-Immunoglobulin Fusion Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,366
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,693
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 30659
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-885-366-17

Query Match 100.0%; Score 12; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 14

US-09-223-342-8
; Sequence 8, Application US/09223342
; Patent No. 6277821
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; APPLICANT: Ammons, William Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Dimer Products
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/223,342
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,504
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-223-342-8

Query Match 100.0%; Score 12; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 15

PCT-US93-04754-17
; Sequence 17, Application PC/TUS9304754
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Grinna, Lynn S
; APPLICANT: Horwitz, Arnold
; TITLE OF INVENTION: BPI-immunoglobulin Fusion Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US93/04754
; FILING DATE: 19930519
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 30659
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US93-04754-17

Query Match 100.0%; Score 12; DB 5; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

Search completed: January 17, 2002, 11:54:39
Job time: 17406 sec

THIS PAGE BLANK (USPCL,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2002, 11:49:53 ; Search time 10436.3 Seconds
(without alignments)
18,969 Million cell updates/sec

Title: US-09-242-202a-22_COPY_1_12
Perfect score: 12
Sequence: 1 GCCACCATGCC 12

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_om.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pl.*
25: em_ro.*
26: em_sts.*
27: em_sy.*
28: em_un.*
29: em_vl.*
30: em_htgo_hum.*
31: em_htgo_inv.*
32: em_htgo_rod.*
33: em_htg_hum.*
34: em_htg_inv.*
35: em_htg_rod.*
36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	12	100.0	22	6	AX139440	AX139440 Sequence
2	12	100.0	23	6	AX098902	AX098902 Sequence
3	12	100.0	23	6	AX099322	AX099322 Sequence
4	12	100.0	27	6	AR027668	AR027668 Sequence
5	12	100.0	27	6	AR050743	AR050743 Sequence
6	12	100.0	27	6	AX030652	AX030652 Sequence
7	12	100.0	27	6	II3573	II3573 Sequence 17
8	12	100.0	27	6	II4327	II4327 Sequence 8
9	12	100.0	27	6	IS0683	IS0683 Sequence 17
10	12	100.0	27	6	I68187	I68187 Sequence 8
11	12	100.0	27	6	I87071	I87071 Sequence 8
12	12	100.0	28	6	I62302	I62302 Sequence 13
13	12	100.0	30	6	AX179323	AX179323 Sequence
14	12	100.0	32	6	AR146549	AR146549 Sequence
15	12	100.0	33	6	AR034283	AR034283 Sequence
16	12	100.0	33	6	AR035389	AR035389 Sequence
17	12	100.0	33	6	AR050804	AR050804 Sequence
18	12	100.0	33	6	AR053809	AR053809 Sequence
19	12	100.0	33	6	AR091590	AR091590 Sequence
20	12	100.0	33	6	AR117469	AR117469 Sequence
21	12	100.0	34	6	AR141722	AR141722 Sequence
22	12	100.0	34	6	AX103386	AX103386 Sequence
23	12	100.0	34	6	I84665	I84665 Sequence 6
24	12	100.0	35	6	AX049981	AX049981 Sequence
25	12	100.0	35	6	AX077144	AX077144 Sequence
26	12	100.0	35	6	AX099890	AX099890 Sequence
27	12	100.0	35	6	AX138011	AX138011 Sequence
28	12	100.0	37	6	AX114583	AX114583 Sequence
29	12	100.0	38	6	AX088738	AX088738 Sequence
30	12	100.0	39	6	AX179597	AX179597 Sequence
31	12	100.0	40	6	AR053605	AR053605 Sequence
32	12	100.0	40	6	AR053608	AR053608 Sequence
33	12	100.0	43	6	AR131268	AR131268 Sequence
34	12	100.0	44	6	AR040911	AR040911 Sequence
35	12	100.0	44	6	AR040950	AR040950 Sequence
36	12	100.0	44	6	AR051818	AR051818 Sequence
37	12	100.0	44	6	AR051857	AR051857 Sequence
38	12	100.0	44	6	AR115954	AR115954 Sequence
39	12	100.0	44	6	AR115993	AR115993 Sequence
40	12	100.0	44	6	AR127321	AR127321 Sequence
41	12	100.0	44	6	AR127360	AR127360 Sequence
42	12	100.0	44	6	I36208	I36208 Sequence 10
43	12	100.0	48	6	A97024	A97024 Sequence 1
44	12	100.0	48	6	AX133393	AX133393 Sequence
45	12	100.0	56	6	AR040909	AR040909 Sequence

ALIGNMENTS

RESULT	1					
AX139440	AX139440	22 bp	DNA			30-MAY-2001
LOCUS	Sequence 3 from Patent EP1094109.					
DEFINITION	AX139440					
ACCESSION	AX139440.1	GI:14275086				
VERSION	human.					
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
TITLE	1 (bases 1 to 22)					
JOURNAL	Human g protein-coupled receptor, pfi-010					
FEATURES	Patent: EP 1094109-A 3 25-APR-2001;					
source	Pfizer Limited (GB) ; PFIZER INC. (US)					
	Location/Qualifiers					
	1..22					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					

BASE COUNT 6 a 9 c 5 g 2 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 5 GCCACCATGGCC 16

RESULT 2

AX098902 LOCUS AX098902 23 bp DNA PAT 02-APR-2001
DEFINITION Sequence 27 from Patent WO0119870.
ACCESSION AX098902
VERSION AX098902.1 GI:13538128
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 23)
AUTHORS Brown,J.P. and Bertelli,F.
TITLE Secreted soluble _g(a)2-g(d)-2, _g(a)2-g(d)-3 or _g(a)2-g(d)-4
calcium channel subunit polypeptides and screening assays using
same

JOURNAL Patent: WO 0119870-A 27 22-MAR-2001;
WARNER-LAMBERT COMPANY (US)

FEATURES
source Location/Qualifiers
1..23
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"

BASE COUNT 2 a 10 c 9 g 2 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 3 GCCACCATGGCC 14

RESULT 3

AX099322 LOCUS AX099322 23 bp DNA PAT 02-APR-2001
DEFINITION Sequence 27 from Patent WO0120336.
ACCESSION AX099322
VERSION AX099322.1 GI:13538474
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 23)
AUTHORS Bertelli,F., Brown,J.P., Dissanayake,V., Suman-Chauhan,N. and
Gee,N.S.

TITLE Screening for alpha2delta-1 subunit binding ligands
JOURNAL Patent: WO 0120336-A 27 22-MAR-2001;
WARNER-LAMBERT COMPANY (US)

FEATURES
source Location/Qualifiers
1..23
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"

BASE COUNT 2 a 10 c 9 g 2 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 3 GCCACCATGGCC 14

RESULT 4

AR027668 LOCUS AR027668 27 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5856302.
ACCESSION AR027668
VERSION AR027668.1 GI:5938488
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Ammons,W.Steve and Little,R.G.
TITLE Therapeutic uses of bactericidal/permeability-increasing protein
dimer products

JOURNAL Patent: US 5856302-A 8 05-JAN-1999;
FEATURES Location/Qualifiers
1..27
/organism="unknown"

BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 5

AR050743 LOCUS AR050743 27 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5827816.
ACCESSION AR050743
VERSION AR050743.1 GI:5973468
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Theofan,G., Horwitz,A., Burke,D., Baltaian,M. and Grinna,L.
TITLE Stable bactericidal/permeability-increasing protein products and
pharmaceutical compositions containing the same

JOURNAL Patent: US 5827816-A 8 27-OCT-1998;
FEATURES Location/Qualifiers
1..27
/organism="unknown"

BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 6

AX030652

LOCUS AX030652 27 bp DNA PAT 20-SEP-2000
DEFINITION Sequence 8 from Patent EP1013760.
ACCESSION AX030652
VERSION AX030652.1 GI:10278178
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Burke,D., Grinna,L., Beltaian,M., Horwitz,A. and Theofan,G.
TITLE Stable bactericidal/permeability-increasing protein products and
pharmaceutical compositions containing the same
JOURNAL Patent: EP 1013760-A 8 28-JUN-2000;
XOMA TECHNOLOGY LT (BM)
FEATURES Location/Qualifiers
source 1..27
/organism="unidentified"
/db_xref="taxon:32644" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 7
LOCUS IL13573 27 bp DNA PAT 26-SEP-1995
DEFINITION Sequence 17 from patent US 5439807.
ACCESSION IL13573
VERSION IL13573.1 GI:996640
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Grinna,L.S.
TITLE Methods for the preparation of endotoxin-binding proteins
JOURNAL Patent: US 5439807-A 17 08-AUG-1995;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 8
LOCUS IL14327 27 bp DNA PAT 26-SEP-1995
DEFINITION Sequence 8 from patent US 5447913.
ACCESSION IL14327
VERSION IL14327.1 GI:997342
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Ammons,W.S. and Little,R.G.

TITLE Therapeutic uses of bactericidal/permeability-increasing protein
dimer products
JOURNAL Patent: US 5447913-A 8 05-SEP-1995;
FEATURES Location/Qualifiers
source 1..27
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 9
LOCUS IL50683 27 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 17 from patent US 5643570.
ACCESSION IL50683
VERSION IL50683.1 GI:2472386
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Theofan,G., Grinna,L.S. and Horwitz,A.
TITLE BPI-immunoglobulin fusion proteins
JOURNAL Patent: US 5643570-A 17 01-JUL-1997;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 10
LOCUS IL68187 27 bp DNA PAT 04-FEB-1998
DEFINITION Sequence 8 from patent US 5674834.
ACCESSION IL68187
VERSION IL68187.1 GI:2830309
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Theofan,G., Horwitz,A., Burke,D., Beltaian,M. and Grinna,L.
TITLE Stable bactericidal/permeability-increasing protein products and
pharmaceutical compositions containing the same
JOURNAL Patent: US 5674834-A 8 07-OCT-1997;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATTGCC 12
|||||
Db 10 GCCACCATTGCC 21

RESULT 11
LOCUS 187071 27 bp DNA 10-JUN-1998
DEFINITION Sequence 8 from patent US 5703038.
ACCESSION 187071
VERSION 187071.1 GI:3206789
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Ammons,W.Steve and Little,R.G.
TITLE Therapeutic uses of bactericidal-permeability-increasing protein dimer products
JOURNAL Patent: US 5703038-A 8 30-DEC-1997;
FEATURES Location/Qualifiers
source
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATTGCC 12
|||||
Db 10 GCCACCATTGCC 21

RESULT 12
LOCUS 162302 28 bp DNA 07-OCT-1997
DEFINITION Sequence 13 from patent US 5658784.
ACCESSION 162302
VERSION 162302.1 GI:2480250
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 28)
AUTHORS Eckner,R., Ewen,M. and Livingston,D.
TITLE Nucleic acid encoding transcription factor p300 and uses of p300
JOURNAL Patent: US 5658784-A 13 19-AUG-1997;
FEATURES Location/Qualifiers
source
BASE COUNT 5 a 13 c 6 g 4 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATTGCC 12
|||||
Db 11 GCCACCATTGCC 22

RESULT 13
LOCUS AX179323 30 bp DNA 03-JUL-2001
DEFINITION Sequence 24 from Patent W00127277.
ACCESSION AX179323

VERSION AX179323.1 GI:14598994
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 30)
AUTHORS Shimkets,R.A., Lichenstein,H. and Boldog,F.L.
TITLE Proteins and polynucleotides encoded thereby
JOURNAL Patent: WO 0127277-A 24 19-APR-2001;
FEATURES Location/Qualifiers
source
1..30
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer 10354784 Rev"
BASE COUNT 5 a 13 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATTGCC 12
|||||
Db 11 GCCACCATTGCC 22

RESULT 14
LOCUS ARI46549 32 bp DNA 08-AUG-2001
DEFINITION Sequence 9 from patent US 6218514.
ACCESSION ARI46549
VERSION ARI46549.1 GI:15109738
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Trikha,M. and Honn,K.V.
TITLE Antibodies specific for soluble truncated integrins
JOURNAL Patent: US 6218514-A 9 17-APR-2001;
FEATURES Location/Qualifiers
source
1..32
/organism="unknown"
BASE COUNT 6 a 10 c 9 g 7 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATTGCC 12
|||||
Db 10 GCCACCATTGCC 21

RESULT 15
LOCUS AR034283 33 bp DNA 29-SEP-1999
DEFINITION Sequence 4 from patent US 5869337.
ACCESSION AR034283
VERSION AR034283.1 GI:5949888
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 33)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and Belshaw,P.
TITLE Regulated transcription of targeted genes and other biological events

JOURNAL Patent: US 5869337-A 4 09-FEB-1999;

FEATURES

Location/Qualifiers

Source

1. .33

/organism="unknown"

BASE COUNT 8 a 13 c 9 g 3 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 12; DB 6; Length 33;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12

|||||

Db 11 GCCACCATGGCC 22

Search completed: January 17, 2002, 11:49:54

Job time: 17206 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:52:06 ; Search time 221.34 Seconds
(without alignments)
11.255 Million cell updates/sec

Title: US-09-242-202a-10_COPY_1_11
Perfect score: 11
Sequence: 1 GCCTTAAGGC 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	11	100.0	17	4	US-09-384-305-17	Sequence 17, Appl
2	11	100.0	17	4	US-09-258-377-22	Sequence 22, Appl
3	11	100.0	33	2	US-08-873-479-39	Sequence 39, Appl
4	11	100.0	33	4	US-09-384-305-23	Sequence 23, Appl
5	11	100.0	33	4	US-09-258-377-28	Sequence 28, Appl
6	11	100.0	33	4	US-09-258-377-31	Sequence 31, Appl
7	11	100.0	34	2	US-08-873-479-19	Sequence 19, Appl
8	11	100.0	34	2	US-08-972-661A-19	Sequence 19, Appl
c 9	11	100.0	36	2	US-08-873-479-18	Sequence 18, Appl
c 10	11	100.0	36	2	US-08-972-661A-18	Sequence 18, Appl
11	11	100.0	37	2	US-09-031-442A-7	Sequence 7, Appl
12	11	100.0	37	4	US-09-258-377-7	Sequence 7, Appl
c 13	11	100.0	40	2	US-08-873-479-46	Sequence 46, Appl
c 14	11	100.0	40	2	US-08-972-661A-32	Sequence 32, Appl
15	11	100.0	42	2	US-08-873-479-37	Sequence 37, Appl
16	11	100.0	42	2	US-09-031-442A-9	Sequence 9, Appl
17	11	100.0	42	2	US-08-972-661A-25	Sequence 25, Appl
18	11	100.0	42	4	US-09-258-377-9	Sequence 9, Appl
19	11	100.0	48	2	US-08-873-479-45	Sequence 45, Appl
20	11	100.0	48	2	US-08-972-661A-31	Sequence 31, Appl
21	11	100.0	54	2	US-09-031-442A-1	Sequence 1, Appl
22	11	100.0	54	4	US-09-384-305-5	Sequence 5, Appl
23	11	100.0	54	4	US-09-258-377-1	Sequence 1, Appl
c 24	11	100.0	57	2	US-08-972-661A-28	Sequence 28, Appl
25	11	100.0	58	2	US-09-031-442A-3	Sequence 3, Appl
26	11	100.0	58	2	US-08-972-661A-27	Sequence 27, Appl
27	11	100.0	58	4	US-09-384-305-7	Sequence 7, Appl

28	11	100.0	58	4	US-09-258-377-3	Sequence 3, Appl
29	11	100.0	185	4	US-09-384-305-20	Sequence 20, Appl
30	11	100.0	185	4	US-09-384-305-21	Sequence 21, Appl
31	11	100.0	185	4	US-09-384-305-22	Sequence 22, Appl
32	11	100.0	185	4	US-09-258-377-25	Sequence 25, Appl
33	11	100.0	185	4	US-09-258-377-26	Sequence 26, Appl
34	11	100.0	185	4	US-09-258-377-27	Sequence 27, Appl
c 35	11	100.0	687	3	US-09-048-052-5	Sequence 5, Appl
c 36	11	100.0	687	4	US-09-160-246-5	Sequence 5, Appl
37	11	100.0	860	1	US-08-310-416A-19	Sequence 19, Appl
38	11	100.0	860	2	US-08-888-171-19	Sequence 19, Appl
39	11	100.0	969	1	US-08-310-416A-12	Sequence 12, Appl
40	11	100.0	969	2	US-08-888-171-12	Sequence 12, Appl
c 41	11	100.0	985	4	US-09-160-246-9	Sequence 9, Appl
c 42	11	100.0	985	4	US-09-160-246-13	Sequence 13, Appl
43	11	100.0	1426	4	US-09-230-380-4	Sequence 4, Appl
44	11	100.0	1878	2	US-08-435-149-17	Sequence 17, Appl
45	11	100.0	1886	6	5256558-1	Patent No. 5256558

ALIGNMENTS

RESULT 1
US-09-384-305-17
; Sequence 17, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952,000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-17

Query Match 100.0%; Score 11; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
|||||
Db 2 gccttaagggc 12

RESULT 2
US-09-258-377-22
; Sequence 22, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Michael D.
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER FILING DATE: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 17
; TYPE: DNA

; ORGANISM: Bacillus
US-09-258-377-22

Query Match 100.0%; Score 11; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952,000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-23

Query Match 100.0%; Score 11; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 3
US-08-873-479-39
; Sequence 39, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 5891701 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174

US-09-258-377-28
; Sequence 28, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-28

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Agtis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Query Match 100.0%; Score 11; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

US-08-873-479-39

RESULT 6
US-09-258-377-31
; Sequence 31, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31

Query Match 100.0%; Score 11; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 5 GCCITTAAGGCG 15

RESULT 4
US-09-384-305-23
; Sequence 23, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin

; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-31

Query Match 100.0%; Score 11; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 2 gccttaaggc 12
|||||

RESULT 7

US-08-873-479-19
; Sequence 19, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agiris, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-19

Query Match 100.0%; Score 11; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 3 GCCTTAAGGC 13
|||||

RESULT 8

US-08-972-661A-19
; Sequence 19, Application US/08872661A
; Patent No. 5958728
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan

; APPLICANT: Sternberg, David
; APPLICANT: Adams, Lee F.
; APPLICANT: Brown, Stephen
; TITLE OF INVENTION: Method For Producing Polypeptides
; TITLE OF INVENTION: In Mutants Of Bacillus Cells
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958728o No. 5958728disk Of No. 5958728th America, Inc.
; STREET: 405 Lexington Avenue - 64th Fl.
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,661A
; FILING DATE: 18-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Staines, Robert L
; REGISTRATION NUMBER: 41,324
; REFERENCE/DOCKET NUMBER: 5111.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-9652
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-972-661A-19

Query Match 100.0%; Score 11; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 3 GCCTTAAGGC 13
|||||

RESULT 9

US-08-873-479-18/c
; Sequence 18, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Agriis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-873-479-18

Query Match 100.0%; Score 11; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
| | | | | | | | | |
Db 22 GCCTTAAGGC 12

RESULT 10
US-08-972-661A-18/C
Sequence 18, Application US/08972661A
Patent No. 5958728

GENERAL INFORMATION:
APPLICANT: Sloma, Alan
APPLICANT: Sternberg, David
APPLICANT: Adams, Lee F.
APPLICANT: Brown, Stephen
TITLE OF INVENTION: Method For Producing Polypeptides
TITLE OF INVENTION: In Mutants Of Bacillus Cells
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958728o No. 5958728disk Of No. 5958728th America, Inc.
STREET: 405 Lexington Avenue - 64th Fl.
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,661A
FILING DATE: 18-NOV-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Starnes, Robert L
REGISTRATION NUMBER: 41,324
REFERENCE/DOCKET NUMBER: 5111.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-878-9652
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-972-661A-18

Query Match 100.0%; Score 11; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
| | | | | | | | | |
Db 22 GCCTTAAGGC 12

RESULT 11
US-09-031-442A-7
Sequence 7, Application US/09031442A
Patent No. 5955310

GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A Polypeptide
TITLE OF INVENTION: In A Bacillus Cell
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5955310o No. 5955310disk of No. 5955310th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,442A
FILING DATE: 26-FEB-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5455.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-031-442A-7

Query Match 100.0%; Score 11; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
| | | | | | | | | |
Db 5 GCCTTAAGGC 15

RESULT 12
US-09-258-377-7
Sequence 7, Application US/09258377
Patent No. 6255076

GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A polypeptide In a
TITLE OF INVENTION: Bacillus Cell
FILE REFERENCE: 5455.200-US
CURRENT APPLICATION NUMBER: US/09/258,377
CURRENT FILING DATE: 1999-02-26
EARLIER APPLICATION NUMBER: 09/031,442

EARLIER FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 7
LENGTH: 37
TYPE: DNA
ORGANISM: Bacillus
US-09-258-377-7

Query Match 100.0%; Score 11; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
DB 5 gccttaaggcc 15

RESULT 13
US-08-873-479-46/C
Sequence 46, Application US/08873479
Patent No. 5891701
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Agis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251.000-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-873-479-46

Query Match 100.0%; Score 11; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
DB 38 GCCTTAAGGC 28

RESULT 14
US-08-972-661A-32/C

Sequence 32, Application US/08972651A
Patent No. 5958728
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
APPLICANT: Sternberg, David
APPLICANT: Adams, Lee F.
APPLICANT: Brown, Stephen
TITLE OF INVENTION: Method For Producing Polypeptides
TITLE OF INVENTION: In Mutants Of Bacillus Cells
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958728o No. 5958728disk of No. 5958728th America, Inc.
STREET: 405 Lexington Avenue - 64th Fl.
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,661A
FILING DATE: 18-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Starnes, Robert L
REGISTRATION NUMBER: 41,324
REFERENCE/DOCKET NUMBER: 5111.200-US
TELEPHONE: 212-878-9652
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-972-661A-32

Query Match 100.0%; Score 11; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
DB 38 GCCTTAAGGC 28

RESULT 15
US-08-873-479-37
Sequence 37, Application US/08873479
Patent No. 5891701
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
APPLICANT: Lynne, Christianson
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agiris, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-873-479-37

```

```

Query Match      100.0%; Score 11; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTTAAGGC 11
        |||||
Db       5 GCCTTAAGGC 15

```

Search completed: January 17, 2002, 11:52:06
Job time: 17253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 07:03:08 ; Search time 10436.3 Seconds
(without alignments)
358.829 Million cell updates/sec

Title: US-09-242-202a-10

Perfect score: 227

Sequence: 1 GCCTTAAGGCCCATATGGTG.....GGTGTCTCTATATATATAT 227

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_om.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_sy.*

28: em_un.*

29: em_vl.*

30: em_htgo_hum.*

31: em_htgo_inv.*

32: em_htgo_rod.*

33: em_htg_hum.*

34: em_htg_inv.*

35: em_htg_rod.*

36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	105.4	46.4	1964	9	HSGROW2	V00520 Human germ
2	105.4	46.4	2771	6	I41411	I41411 Sequence 5
3	105.4	46.4	4456	6	AR119909	AR119909 Sequence
4	105.4	46.4	4874	12	AF239248	AF239248 Eukaryoti
5	98	43.2	1827	9	AF237583	AF237583 Homo sapi
6	96	42.3	831	6	I02857	I02857 Sequence 3
7	96	42.3	2086	6	AR102885	AR102885 Sequence
8	96	42.3	2086	6	AR108064	AR108064 Sequence
9	96	42.3	2086	6	AR134712	AR134712 Sequence
10	96	42.3	2657	9	HUMGHN	M13438 Human growt
11	96	42.3	3141	6	A18662	A18662 Synthetic c
12	96	42.3	3141	6	A18663	A18663 Synthetic c
13	96	42.3	3141	6	A20119	A20119 Hybrid insu
14	96	42.3	3141	6	A20120	A20120 Hybrid insu
15	96	42.3	4495	12	AF369966	AF369966 Cloning v
16	96	42.3	4650	12	AF396260	AF396260 Cloning v
17	96	42.3	4657	12	AF239249	AF239249 Eukaryoti
18	96	42.3	4665	12	AF239250	AF239250 Eukaryoti
19	96	42.3	4707	12	AF067196	AF067196 Cloning v
20	96	42.3	4713	12	AF067197	AF067197 Cloning v
21	96	42.3	4724	12	AF239247	AF239247 Eukaryoti
22	96	42.3	4775	12	AF239251	AF239251 Eukaryoti
23	96	42.3	4912	12	AF239252	AF239252 Eukaryoti
24	96	42.3	5646	9	HSA7670	AJ007670 Homo sapi
25	96	42.3	9274	6	AX060703	AX060703 Sequence
26	96	42.3	66495	9	HUMGHCSA	J03071 Human growt
27	96	42.3	159128	2	AC040958	AC040958 Homo sapi
28	96	42.3	159387	2	AC084858	AC084858 Homo sapi
29	95.6	42.1	1944	9	AF237585	AF237585 Homo sapi
30	95	41.9	6776	12	PVPHM3	X76682 Plasmid vec
31	95	41.9	8451	12	PVPHM3BL	X76683 Plasmid vec
32	87	38.3	2660	6	E00140	E00140 Genomic DNA
33	87	38.3	2660	9	HUMGHV	K00470 Homo sapien
34	87	38.3	2740	9	HUMCS3	M15894 Human chori
35	87	38.3	159128	2	AC040958	AC040958 Homo sapi
36	87	38.3	191583	9	AC015651	AC015651 Homo sapi
37	86.4	38.1	1125	6	I02858	I02858 Sequence 4
38	86.4	38.1	2033	9	AF237586	AF237586 Homo sapi
39	86.4	38.1	2967	9	HUMPLA	J00289 Homo sapien
40	85.4	37.6	2594	9	MM002293	U02293 Macaca mula
41	75	33.0	2193	9	AF237584	AF237584 Homo sapi
42	74.2	32.7	6160	1	ECT7DN3	Z32693 E.coli pT7h
43	74.2	32.7	11271	1	ECT7DN1	Z32691 E.coli (HMI
44	74	32.6	642	9	AF110644	AF110644 Homo sapi
45	74	32.6	796	6	E00009	E00009 mRNA coding

ALIGNMENTS

RESULT 1

HSGROW2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

2 (bases 1 to 1964)

AUTHORS

TITLE

JOURNAL

MEDLINE

2 (bases 1 to 1964)

AUTHORS

TITLE

HSGROW2 1964 bp DNA PRI 10-FEB-1999
Human germ line gene for growth hormone (presomatotropin).

V00520 J00148 K00612

V00520.1 GI:31906

complementary DNA; germ line; growth hormone; signal peptide.
human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

1 (bases 1 to 1964)

DeNoto,F.M., Moore,D.D. and Goodman,H.M.

Human growth hormone DNA sequence and mRNA structure: possible

alternative splicing

Nucleic Acids Res. 9 (15), 3719-3730 (1981)

82014939

2 (bases 1 to 1964)

Adelman,J.P., Hayflick,J.S., Vasser,M. and Seeburg,P.H.

In vitro deletional mutagenesis for bacterial production of the

20,000-dalton form of human pituitary growth hormone

DNA 2 (3), 183-193 (1983)

84057143

This entry was previously called <HSGROWL>.

See <HSGROWL> for mRNA sequence (with some differences).

Location/Qualifiers

1..1964

/organism="Homo sapiens"

/db_xref="taxon:9606"

(274,275)..1905

join(274,275)..344,601..761,971..1090,1184..1348,

1602..1905)

(274,275)..344

/number=1..344

join(335..344,601..761,971..1090,1184..1348,1602..1799)

/codon_start=1

/product="growth hormone"

/protein_id="CAA23779.1"

/db_xref="GI:312406"

/translation="MATGSRSTLLAFGLLCPWLQSGAFPTIPLSLFDNASLRAH

RLHQAEDTTQEEFEATIPKEQKYSFLQNTSLCFSESPTPSNRETOQKSNLELL

RISLLLSWLEPWFQFVRSFVANSVLVYSGDSNVYDLKDLLEGIQITLMGRLEDGSPR

TGQIFKQTSKEDTNSHNDALLKNYGLLYCFKRDMDKVETFLRIVQCRSVEGSGF"

345..600

/number=1

601..761

/number=2

762..970

/number=2

971..1090

/number=3

1091..1183

/number=3

1184..1348

/number=4

1349..1601

/number=4

1602..1905

/number=5

BASE COUNT 451 a 550 c 550 g 413 t

ORIGIN

Query Match 46.4%; Score 105.4; DB 9; Length 1964;

Best Local Similarity 99.1%; Pred. No. 7.6e-21;

Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 121 AGTGCCTCTCCTGGCCCTGGAAAGTTGCCACTCCAGTGGCCACCGCCTGTCTTAATAAA 180

|||||

Db 1832 AGTGCCTCTCCTGGCCCTGGAAAGTTGCCACTCCAGTGGCCACCGCCTGTCTTAATAAA 1891

|||||

Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATAATATAT 227

|||||

Db 1892 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATAATATAT 1938

|||||

RESULT 2

I41411

LOCUS I41411 2771 bp DNA PAT 13-MAY-1997

DEFINITION Sequence 5 from patent US 5625124.

ACCESSION I41411

VERSION I41411.1 GI:2082001

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2771):

FAUK, P. and Gordon, J. I.

Animal model for helicobacter pylori infection

JOURNAL Patent: US 5625124-A 5 29-APR-1997;

Location/Qualifiers

1..2771

source

BASE COUNT 666 a 718 c 704 g 683 t

ORIGIN

Query Match 46.4%; Score 105.4; DB 6; Length 2771;

Best Local Similarity 99.1%; Pred. No. 7.5e-21;

Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 121 AGTGCCTCTCCTGGCCCTGGAAAGTTGCCACTCCAGTGGCCACCGCCTGTCTTAATAAA 180

|||||

Db 2177 AGTGCCTCTCCTGGCCCTGGAAAGTTGCCACTCCAGTGGCCACCGCCTGTCTTAATAAA 2236

|||||

Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATAATATAT 227

|||||

Db 2237 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATAATATAT 2283

|||||

RESULT 3

AR119909

LOCUS AR119909 4456 bp DNA PAT 16-MAY-2001

DEFINITION Sequence 1 from patent US 6153427.

ACCESSION AR119909

VERSION AR119909.1 GI:14102608

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4456)

KING, D. and Schneider, G. B.

Erythropoietin-inducible, erythroid-specific DNA construct

JOURNAL Patent: US 6153427-A 1 28-NOV-2000;

Location/Qualifiers

1..4456

source

BASE COUNT 1238 a 978 c 1076 g 1164 t

ORIGIN

Query Match 46.4%; Score 105.4; DB 6; Length 4456;

Best Local Similarity 99.1%; Pred. No. 7.3e-21;

Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 121 AGTGCCTCTCCTGGCCCTGGAAAGTTGCCACTCCAGTGGCCACCGCCTGTCTTAATAAA 180

|||||

Db 4324 AGTGCCTCTCCTGGCCCTGGAAAGTTGCCACTCCAGTGGCCACCGCCTGTCTTAATAAA 4383

|||||

Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATAATATAT 227

|||||

Db 4384 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATAATATAT 4430

|||||

RESULT 4

AF239248

LOCUS AF239248 4874 bp DNA circular SYN 12-APR-2000

DEFINITION Eukaryotic expression vector pCMV4, complete sequence.

ACCESSION AF239248

VERSION AF239248.1 GI:7542545

KEYWORDS

SOURCE Expression vector pCMV4.

ORGANISM Expression vector pCMV4.

REFERENCE 1 (bases 1 to 4874)

ANDERSSON, S., DAVIS, D. L., DAHLBACK, H., JORNVAALL, H. and RUSSELL, D. W.

Cloning, structure, and expression of the mitochondrial cytochrome

P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme

JOURNAL J. Biol. Chem. 264 (14), 8222-8229 (1989)

MEDLINE 89255259

PUBMED 2722778

REFERENCE 2 (bases 1 to 4874)

ANDERSSON, S., DAVIS, D. L., DAHLBACK, H., JORNVAALL, H. and RUSSELL, D. W.

Direct Submission

JOURNAL Submitted (25-FEB-2000) Molecular Genetics, University of Texas

Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd.,
Dallas, TX 75235, USA

FEATURES
source Location/Qualifiers
1. .4874

/organism="Expression vector pCMV4"

/db_xref="taxon:120119"

BASE COUNT 1165 a 1237 c 1170 g 1302 t

ORIGIN

Query Match 46.4%; Score 105.4; DB 12; Length 4874;
Best Local Similarity 99.1%; Pred. No. 7.3e-21;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCTTAATAAA 180

Db 1055 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCTTAATAAA 1114

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCCTCTCTATAATATAT 227

Db 1115 ATTAAGTTGCATCATTTTGTCTGACTAGTGTCCTCTCTATAATATAT 1161

RESULT 5
LOCUS AF237583 1827 bp DNA PRI 11-MAY-2001
DEFINITION Homo sapiens recombinant IgG1 heavy chain gene, partial cds.
ACCESSION AF237583
VERSION AF237583.1 GI:9857752

KEYWORDS human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1827)
AUTHORS Vidarsson, G., Jansen, M., Boel, E. and van de Winkel, J.G.J.
TITLE Activity of Human IgG and IgA Subclasses in Immune Defense Against
Neisseria meningitidis Serogroup B
J. Immunol. 166 (10), 6250-6256 (2001)

2 (bases 1 to 1827)
AUTHORS Vidarsson, G., Jansen, M., Boel, E. and van de Winkel, J.G.J.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2000) Department of Immunology, University
Medical Center Utrecht, Rm. KC.02-085.2, Lundlaan 6, Utrecht 3584
EA, The Netherlands

FEATURES
source Location/Qualifiers
1. .1827
/organism="Homo sapiens"
/db_xref="taxon:9606"

1. .294
/note="CH1"
join(<1. .294, 686. .730, 849. .1178, 1276. .>1598)
/product="recombinant IgG1 heavy chain"
join(<1. .294, 686. .730, 849. .1178, 1276. .1598)
/codon_start=3
/product="recombinant IgG1 heavy chain"
/protein_id="AAG00909.1"
/db_xref="GI:9857753"

/translation="STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGA
LTSGVHPFPAVLOSGGLYSSIVYFPPSGLIGTQTYICNVNHKPSNTKYDKRVEPKSC
DKTCTCPAPELGGPSVFLPPKPKDTLMISITPELVTCVVVDVSHEDPEVKFNWY
VDGVEHNKPTPREQYNSTIRYSIVLTVLHODWLNKEYKKVSKALPAPTEKTI
SKAGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTT
PPVLDSDGSFFLYSKLVDKSRWQQGVFSCVMHEALHNHYTQKSLSLSPGK"

686. .730
/note="hinge"
849. .1178
/note="CH2"
1276. .1598
/note="CH3"

exon
mRNA
CDS

exon
exon
exon

BASE COUNT 390 a 621 c 498 g 316 t 2 others
ORIGIN

Query Match 43.2%; Score 98; DB 9; Length 1827;
Best Local Similarity 78.3%; Pred. No. 1.2e-18;
Matches 130; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 63 TCAGAGCCCCGGGCGACGACAGGCCAATGCCCGTCTTCCCTGCAGGATGAGTAGAG 122

Db 1578 TCCCTGTCCCCGGGTAATAGTAGCGGTACCGGTGGCATCCCTGTGACCCCTCCCCAG 1637

QY 123 TGCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCTTGTCCTAATAAAAT 182

Db 1638 TGCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCTTGTCCTAATAAAAT 1697

QY 183 TAAAGTTGCATCATTTTGTCTGACTAGGTGTCC-TCTATAATAATAT 227

Db 1698 TAAAGTTGCATCATTTTGTCTGACTAGGTGTCC-TCTATAATAATAT 1743

RESULT 6
LOCUS I02857 831 bp ss-DNA PAT 18-MAY-1993
DEFINITION Sequence 3 from Patent US 4446235.
ACCESSION I02857
VERSION I02857.1 GI:267919

KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 831)
AUTHORS Seeburg, P.H.
TITLE Method for cloning human growth hormone variant genes
JOURNAL Patent: US 4446235-A 3 01-MAY-1984;
Genentech, Inc.;
South San Francisco, CA

FEATURES
source Location/Qualifiers
1. .831
/organism="unknown"
BASE COUNT 175 a 231 c 219 g 206 t
ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 831;
Best Local Similarity 99.1%; Pred. No. 5e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCTTGTCCTAATAAA 180

Db 231 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCTTGTCCTAATAAA 290

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCC-TCTATAATAATAT 227

Db 291 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCC-TCTATAATAATAT 338

RESULT 7
LOCUS AR102885 2086 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 9 from patent US 6087129.
ACCESSION AR102885
VERSION AR102885.1 GI:12814473

KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2086)
AUTHORS Newgard, C.B., Normington, K.D., Clark, S.A., Thigpen, A.E., Quaaide, C.
and Kruse, F.
TITLE Recombinant expression of proteins from secretory cell lines
JOURNAL Patent: US 6087129-A 9 11-JUL-2000;
FEATURES Location/Qualifiers

```

source
BASE COUNT 456 a 604 c 560 g 466 t
ORIGIN
Query Match 42.3%; Score 96; DB 6; Length 2086;
Best Local Similarity 99.1%; Pred. No. 4.8e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCACCCAGCCCTGTCTTAATAAA 180
Db 1559 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCACCCAGCCCTGTCTTAATAAA 1618
QY 181 ATTAAGTTCATCATTTTGTCTGACTAGGTGTC-TCTATAATATAT 227
Db 1619 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATAT 1666

RESULT 8
LOCUS ARI08064 2086 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 9 from patent US 6110707.
ACCESSION ARI08064
VERSION ARI08064.1 GI:12823551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2086)
AUTHORS Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
Quade,C., Kruse,F. and McGarry,D.
TITLE Recombinant expression of proteins from secretory cell lines
JOURNAL Patent: US 6110707-A 9 29-AUG-2000;
MEDLINE Location/Qualifiers
COMMENT 1. .2086
FEATURES
source
BASE COUNT 456 a 604 c 560 g 466 t
ORIGIN
Query Match 42.3%; Score 96; DB 6; Length 2086;
Best Local Similarity 99.1%; Pred. No. 4.8e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCACCCAGCCCTGTCTTAATAAA 180
Db 1559 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCACCCAGCCCTGTCTTAATAAA 1618
QY 181 ATTAAGTTCATCATTTTGTCTGACTAGGTGTC-TCTATAATATAT 227
Db 1619 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATAT 1666

RESULT 9
LOCUS ARI34712 2086 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 9 from patent US 6194176.
ACCESSION ARI34712
VERSION ARI34712.1 GI:14123617
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2086)
AUTHORS Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
Quade,C. and Kruse,F.
TITLE Recombinant expression of proteins from secretory cell lines
JOURNAL Patent: US 6194176-A 9 27-FEB-2001;
MEDLINE Location/Qualifiers
COMMENT 1. .2086
FEATURES
source
BASE COUNT 456 a 604 c 560 g 466 t
ORIGIN

```

ORIGIN

```

Query Match 42.3%; Score 96; DB 6; Length 2086;
Best Local Similarity 99.1%; Pred. No. 4.8e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCACCCAGCCCTGTCTTAATAAA 180
Db 1559 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCACCCAGCCCTGTCTTAATAAA 1618
QY 181 ATTAAGTTCATCATTTTGTCTGACTAGGTGTC-TCTATAATATAT 227
Db 1619 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATAT 1666

RESULT 10
LOCUS HUMGHN 2657 bp DNA PRI 29-APR-1996
DEFINITION Human growth hormone gene (HGH-N), complete cds.
ACCESSION M13438
VERSION M13438.1 GI:183156
KEYWORDS Alu repeat; growth hormone; hormone; repeat region.
SOURCE Homo sapiens (clone: HGH-N.) (tissue library: Lawn et al.) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2657)
AUTHORS Seeburg,P.H.
TITLE The human growth hormone gene family: nucleotide sequences show
recent divergence and predict a new polypeptide hormone
JOURNAL DNA 1 (3), 239-249 (1982)
MEDLINE 83182010
COMMENT The Alu family sequences are known to be transcribed by RNA
polymerase III and in [1] is inserted such that transcription would
be from the opposite strand to that of the growth hormone genes.
FEATURES
source
1. .2657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HGH-N."
/tissue_lib="Lawn et al."
/map="17q21-qter"
497. .568
/genes="GH-N"
/number=1
prim_transcript 497. .2129
/genes="GH-N"
/note="HGH-N mRNA"
497. .2129
/genes="GH-N"
join(559. .568,828. .895)
sig_peptide
/genes="GH-N"
join(559. .568,828. .988,1196. .1315,1409. .1573,1827. .2024)
CDS
/genes="GH-N"
/note="precursor"
/codon_start=1
/product="growth hormone"
/protein_id="AAA98618.1"
/db_xref="GI:183157"
/translation="MATGSRTSLLAFLGALLCLPLQEGSAFPTPLSLFDNMLRAH
RLHQLAFDTYQEFEEAYIPKEQYSLQNPQSLCSESIPTPSNREETQQRKLELL
RISLLLIQSWLEPQVFLRSVFNLSVYGASDNYDLKLEEGTQTLMLGRLEDSPR
TGQIFQYKSFEDTNSHNDALLKLYGLLYCFRKMDKVFETFLRIVQCRSEVSGGF"
569. .827
/genes="GH-N"
/number=1
828. .988
/genes="GH-N"
/number=2
join(896. .988,1196. .1315,1409. .1573,1827. .2021)
mat_peptide
/genes="GH-N"
/product="growth hormone"

```

```
intron      989..1195
/gene="GH-N"
/number=2
exon        1196..1315
/gene="GH-N"
/number=3
intron      1316..1408
/gene="GH-N"
/number=3
exon        1409..1573
/gene="GH-N"
/number=4
intron      1574..1826
/gene="GH-N"
/number=4
exon        1827..2129
/gene="GH-N"
/number=5
polyA_signal 2111..2116
/gene="GH-N"
repeat_region complement(2228..2501)
/rpt_family="Alu"
BASE COUNT 598 a 741 c 734 g 584 t
ORIGIN      1 bp upstream of EcoRI site.

Query Match      42.3%; Score 96; DB 9; Length 2657;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCTCTCTCGGCCGCGGAGTTGGCACTCAGTGCACCGCCAGCCGTTGTCTATAATAA 180
|||||
DB 2057 AGTGCTCTCTCGGCCGCGGAGTTGGCACTCAGTGCACCGCCAGCCGTTGTCTATAATAA 216
|||||
QY 181 ATTAAGTTGCATCATTTGTCTGACTAGTGTGCC-TCTATAATATAT 227
|||||
DB 2117 ATTAAGTTGCATCATTTGTCTGACTAGTGTCTCTATAATATAT 2164

RESULT 11
A18662
LOCUS      A18662      3141 bp      mRNA      PAT      10-MAY-1994
DEFINITION Synthetic cDNA sequence for hybrid insulin/IGF-I receptor.
ACCESSION A18662
VERSION   A18662.1 GI:513336
KEYWORDS  synthetic construct.
SOURCE    synthetic construct.
ORGANISM  artificial sequence.
REFERENCE 1 (bases 1 to 3141)
AUTHORS   MAMMALIAN CELLS EXPRESSING A HYBRID RECEPTOR
TITLE     Patent: WO 9117253-A 14 14-NOV-1991;
JOURNAL   Location/Qualifiers
FEATURES   source
           1..3141
           /organism="synthetic construct"
           /db_xref="taxon:32630"
           12..2858
           /codon_start=1
           /product="polypeptide of extracellular domain of hybrid
           insulin/IGF-I receptor"
           /protein_id="CAA01405.1"
           /db_xref="GI:513337"
           /translation="MKSGGGGSPSLWGLFLSAAISLWPTSIGICPGIDIRNDYQ
           QLRLENTVEGYLHLLISKAEDYRFPKLTIVITEYLLFRVAGLSGLDFPN
           LTVRWKLFYNYALVIFEMNLKIDGLYNLRNITRGAIRKRNADLCYLSVDSWLSI
           LDVSNVYVGNKPKCGDLCPGTMEKPKTTINNEYVRCWTNNRCQKWCPT
           CGKTRACENNECCHEPCLGSCSAPNDTACVACHYVYAGVCPACPPNTYRFGWRC
           VDRFCANILSAESDSGEGFVHDEGMQECPSGFIIRNSNLLCTPLGCPKVCHELL
           EGKTDIVSTSAQLRGTVINGSLIINIRGNNALEAFLEANLIEISYKIRRS
           YALVSLFFGKRLRIGETLEIGNYSFALDNQNLRLQMDWSKHNLTTCQKLPFHYN
           PKLCLSETHKMEVSGTKGRQERNDIALKTNGDKASCENELLKFSYIRTSFDKILLRW

CDS
Query Match      42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCTCTCTCGGCCGCGGAGTTGGCACTCAGTGCACCGCCAGCCGTTGTCTATAATAA 180
|||||
DB 2898 AGTGCTCTCTCGGCCGCGGAGTTGGCACTCAGTGCACCGCCAGCCGTTGTCTATAATAA 2957
|||||
QY 181 ATTAAGTTGCATCATTTGTCTGACTAGTGTGCC-TCTATAATATAT 227
|||||
DB 2958 ATTAAGTTGCATCATTTGTCTGACTAGTGTCTCTATAATATAT 3005
|||||

RESULT 12
A18663/c
LOCUS      A18663      3141 bp      mRNA      PAT      10-MAY-1994
```

```
EPYWPDPDRLLGFMLFYKEAPYQNVTEFDQDAGCSNSWTVDIDPLRSDPKSQN
HPGWLMRGLKWPQYAIKVTIYVSDERRTYGAKSDIYYVQTDATNPSPVLPDPSVS
NSSQIILKWKPPDPNGNITHYLVFWEQAEDELFDYCLAGLKLPSRTWSPPE
SEDSKUNQSEVEDSAGECCSCPKTDSQILKELESSRKTFEEDYLNHNVFVPRKTS
GTGAEDRPSKRRSLGDNVTVAVPYVAAPNTSSTSVTSPEEHRPFKVKYNKES
LVIISGLRHTGYRIELQACNQDTPFEERCVAAYVSARTMPKAKADDIVGPTHELFEN
NVHLMQEPKEPNGLIVLYEVSRYRGGDEELHLCVSRKHFALEGRCLRGLSPGNY
VRIATSLAGNSWTEPTYFYVTDYLDVPSNIAK"
2859..2888
/partial
/codon_start=1
/transl_table=11
/product="polypeptide of extracellular domain of hybrid
insulin/IGF-I receptor"
/protein_id="CAA01406.1"
/db_xref="GI:583136"
/translation="SRGSGVASL"
2889..2981
/partial
/codon_start=1
/transl_table=11
/product="polypeptide of extracellular domain of hybrid
insulin/IGF-I receptor"
/protein_id="CAA01407.1"
/db_xref="GI:513338"
/translation="PLPSAPSPGSGCHSSAHQPCPNKIKLHFFV"
2982..2999
/partial
/codon_start=1
/transl_table=11
/product="polypeptide of extracellular domain of hybrid
insulin/IGF-I receptor"
/protein_id="CAA01408.1"
/db_xref="GI:583137"
/translation="LGULL"
3000..3059
/partial
/codon_start=1
/transl_table=11
/product="polypeptide of extracellular domain of hybrid
insulin/IGF-I receptor"
/protein_id="CAA01409.1"
/db_xref="GI:513339"
/translation="YVGVGGGMEQGASWEDNL"
3060..3141
/partial
/codon_start=1
/transl_table=11
/product="polypeptide of extracellular domain of hybrid
insulin/IGF-I receptor"
/protein_id="CAA01410.1"
/db_xref="GI:513340"
/translation="GLRGLGTGKLECSGTLIAHCLNRLGGS"
717 a 892 c 863 g 669 t
BASE COUNT
ORIGIN

Query Match      42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCTCTCTCGGCCGCGGAGTTGGCACTCAGTGCACCGCCAGCCGTTGTCTATAATAA 180
|||||
DB 2898 AGTGCTCTCTCGGCCGCGGAGTTGGCACTCAGTGCACCGCCAGCCGTTGTCTATAATAA 2957
|||||
QY 181 ATTAAGTTGCATCATTTGTCTGACTAGTGTGCC-TCTATAATATAT 227
|||||
DB 2958 ATTAAGTTGCATCATTTGTCTGACTAGTGTCTCTATAATATAT 3005
|||||

RESULT 12
A18663/c
LOCUS      A18663      3141 bp      mRNA      PAT      10-MAY-1994
```

DEFINITION Synthetic cDNA sequence for hybrid insulin/IGF-I receptor.
ACCESSION A18663
VERSION A18663.1 GI:512264
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3141)
AUTHORS
TITLE MAMMALIAN CELLS EXPRESSING A HYBRID RECEPTOR
JOURNAL Patent: WO 9117253-A 15 14-NOV-1991;
FEATURES Location/Qualifiers
source
1. .3141
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 669 a 863 c 892 g 717 t
ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 121 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACGAGCCTTGCTCCTAATAAA 180
|||||
Db 244 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACGAGCCTTGCTCCTAATAAA 185
|||||
QY 181 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227
|||||
Db 184 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 137
|||||

RESULT 13
A20119
LOCUS A20119 3141 bp mRNA PAT 08-AUG-1994
DEFINITION Hybrid insulin/IGF-I receptor mRNA.
ACCESSION A20119
VERSION A20119.1 GI:583286
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3141)
AUTHORS
TITLE A HYBRID CELLULAR RECEPTOR
JOURNAL Patent: WO 9117252-A 15 14-NOV-1991;
FEATURES Location/Qualifiers
source
1. .3141
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 717 a 892 c 863 g 669 t
ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 121 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACGAGCCTTGCTCCTAATAAA 180
|||||
Db 2898 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACGAGCCTTGCTCCTAATAAA 2957
|||||
QY 181 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227
|||||
Db 2958 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 3005
|||||

RESULT 14
A20120/c
LOCUS A20120 3141 bp mRNA PAT 08-AUG-1994
DEFINITION Hybrid insulin/IGF-I receptor mRNA.
ACCESSION A20120
VERSION A20120.1 GI:578998

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3141)
AUTHORS
TITLE A HYBRID CELLULAR RECEPTOR
JOURNAL Patent: WO 9117252-A 16 14-NOV-1991;
FEATURES Location/Qualifiers
source
1. .3141
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 669 a 863 c 892 g 717 t
ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 121 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACGAGCCTTGCTCCTAATAAA 180
|||||
Db 244 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACGAGCCTTGCTCCTAATAAA 185
|||||
QY 181 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227
|||||
Db 184 ATTAAGTTCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 137
|||||

RESULT 15
AF369966
LOCUS AF369966 4495 bp DNA circular SYN 22-JUL-2001
DEFINITION Cloning vector pCMV-MCS, complete sequence.
ACCESSION AF369966
VERSION AF369966.1 GI:14994098
KEYWORDS
SOURCE Cloning vector pCMV-MCS.
ORGANISM Cloning vector pCMV-MCS.
REFERENCE 1 (bases 1 to 4495)
AUTHORS Grafsky,A.J. III.
TITLE pCMV-MCS for use in the AAV Helper-Free System
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4495)
AUTHORS Grafsky,A.J. III.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2001) Technical Services, Stratagene, 11011 N.
Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES Location/Qualifiers
source
1. .4495
/organism="Cloning vector pCMV-MCS"
/db_xref="taxon:161367"
/lab_host="Escherichia coli"
1. .670
/note="CMV promoter"
678. .1170
/note="similar to beta-globin intron"
1178. .1253
/note="MCS"
1249. .1735
/note="hgh/poly A"
complement(1844. .2701)
/note="ColE1 origin"
complement(2711. .3571)
/note="confers ampicillin resistance"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAK76424.1"
/db_xref="GI:14994098"
/translation="MSIOHFRVALIPFFAFCILPFAHETLVKVKDAEDQIGARVGY
IELDLSNGKILSFSPREPEPMMSTFKVLLCGNLSRIDAGQQLGRRIRHYSNDLVE
YSPVTEKHLTDGNTVRELCSAITSNDNTAANLLLTITGGPKELTAFLNMGDHTRL

promoter
misc_feature
misc_feature
misc_feature
rep_origin
CDS

DRWPELNCIAPNDRGTTMPVAMATTIRKLLTGELLTTLASRQOOLDWNEADKVAQGL
LRSALPAGFWFIADKSGAGERSGIIAALGPDGKPSRIVVITYTTSQATMDERNRQIA
EIGASLIKHW ^a
complement(3952..4413)
/note="fl origin"
rep_origin
BASE COUNT
1124 a 1117 c 1101 g 1153 t

[illegible]

Search completed: January 17, 2002, 11:48:13
Job time: 17105 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: January 17, 2002, 07:06:53 ; Search time 599.86 Seconds
(without alignments)
324.430 Million cell updates/sec

Title: US-09-242-202a-10
Perfect score: 227
Sequence: 1 GCCTTAAGGCCATATGTTG.....GGTGTCTCTATAATATTAT 227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227	100.0	227	19	AAV21719
2	214.4	94.4	228	19	AAV21723
3	214.4	94.4	1547	19	AAV21732
4	214.4	94.4	1807	19	AAV21733
5	214.4	94.4	2308	19	AAV21734
6	160.6	70.7	1425	19	AAV21724
7	160.6	70.7	2125	19	AAV21727
8	152	67.0	1911	19	AAV21726
9	105.4	46.4	2771	18	AAV76770
10	105.4	46.4	4456	22	AAC83268
11	96	42.3	191	19	AAV33343

12	96	42.3	191	20	AAZ40410	Human growth hormo
13	96	42.3	191	21	AAZ50393	Human growth hormo
14	96	42.3	563	22	AAF76860	Human secreted pro
15	96	42.3	1441	20	AAI15886	Not1/Sali psk-GHRH
16	96	42.3	2086	21	AAC55720	Secretory cell DNA
17	96	42.3	2086	22	AAF58805	Human growth hormo
18	96	42.3	2160	18	AAT62826	Human growth hormo
19	96	42.3	3141	12	AAQ14381	Extracellular doma
20	96	42.3	3141	12	AAQ14813	Soluble human insu
21	96	42.3	3141	12	AAQ14815	Hybrid human insul
22	96	42.3	3426	20	AAZ40416	Plasmid pIF0836 co
23	96	42.3	3589	20	AAZ40418	Plasmid pIF0921 co
24	96	42.3	3599	19	AAV50428	Plasmid pIG0552 lo
25	96	42.3	3599	19	AAV40796	Actual sequence of
26	96	42.3	3600	19	AAV50427	Plasmid pIG0552 up
27	96	42.3	3600	19	AAV40795	Expected sequence
28	96	42.3	5686	20	AAZ40415	Plasmid pIN1143 co
29	96	42.3	6513	21	AAV5627	Nucleotide sequenc
30	96	42.3	9164	20	AAZ82259	Beta-domain delete
31	96	42.3	9274	22	AAF24297	HER2 transgene pla
32	96	42.3	9534	18	AAT62072	Vector RP3224E2 en
33	96	42.3	11846	20	AAZ82261	Factor VIII protei
34	96	42.3	12022	20	AAZ82260	Factor VIII protei
35	93.4	41.1	3653	18	AAZ61522	Cosmid CV014 conta
36	87	38.3	2660	4	AAZ30032	Sequence of gene f
37	87	38.3	2660	4	AAZ30032	Human cancer speci
38	79	34.8	2994	22	AAF84109	Met-des(F1-P2-P3-I
39	74	32.6	682	15	AAT02620	Met-des(F1-P2-P3-I
40	74	32.6	682	16	AAT11228	Met-des(Phe1-Pro2-
41	74	32.6	700	16	AAT02619	Met-des(Phe1-Pro2-
42	74	32.6	700	16	AAT11227	Met-des(Phe1-Pro2-
43	74	32.6	706	16	AAT02618	Met-des(Phe1-Pro2)
44	74	32.6	706	16	AAT11225	Met-des(Phe1-Pro2)
45	74	32.6	709	16	AAT02617	Met-des(Phe1)-soma

ALIGNMENTS

RESULT 1	
AAV21719	AAV21719 standard; cDNA; 227 BP.
ID	AAV21719 standard; cDNA; 227 BP.
XX	
AC	AAV21719;
XX	
DT	17-AUG-1998 (first entry)
XX	
DE	Combined 3' splice sequence and polyA tail.
XX	
KW	Vector; vaccine; tumour; antigen; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO9806863-A1.
XX	
PD	19-FEB-1998.
XX	
PF	14-AUG-1997; 97WO-US14306.
XX	
PR	14-AUG-1996; 96US-0023931.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Nelson EL, Nelson PJ;
XX	
DR	WPI; 1998-159552/14.
XX	
PT	Humanised polynucleotide vectors - comprising human derived promoter
PT	and sequence acceptance site, used for the production of vaccines
XX	
PS	Disclosure; Page 17; 125pp; English.
XX	

CC This DNA sequence comprises a combined 3' splice sequence and
CC polyA tail sequence derived from human gene sequences. Novel
CC humanised vectors of the invention (see AAV21724, AAV21727 and
CC AAV21732-34) comprise a human-derived promoter or mammalian homologue
CC which is functional in mammalian target tissue and cells and an
CC acceptance site which accepts cDNA products from RT-PCR cloning.
CC They may also include human-derived splice-polyA sequences. The
CC vectors are used to express target antigens, especially tumour
CC antigens. They are non-replicating in mammalian cells but are
CC capable of extended stable expression of target sequences generating
CC an immune response in immunised individuals. The vectors
CC selectively elicit immune responses to the target sequences with
CC little or no immune response to the other components of the vectors.
CC The target antigens are expressed as intracellular polypeptides or
CC peptides and, as such, are processed as self polypeptides or
CC peptides and appropriately presented on antigen presenting cells.
XX
SQ Sequence 227 BP; 47 A; 64 C; 60 G; 56 T; 0 other;

Query Match 100.0%; Score 227; DB 19; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.4e-62;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGGATGGGAGACCTGT 60
DB 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGGATGGGAGACCTGT 60
QY 61 AGTCAGAGCCCCGGGCGAGCAGCCCAATGCCGCTCTCCCTGCAGGATGAGTAGTG 120
DB 61 AGTCAGAGCCCCGGGCGAGCAGCCCAATGCCGCTCTCCCTGCAGGATGAGTAGTG 120
QY 121 AGTCGCTCTCTCTGCGCCCTGGGAAGTTGCCACATCCAGTGCCACCAGCCCTTGTCCTTAATAA 180
DB 121 AGTCGCTCTCTGCGCCCTGGGAAGTTGCCACATCCAGTGCCACCAGCCCTTGTCCTTAATAA 180
QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227
DB 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227

RESULT 2
AAV21723
ID AAV21723 standard; cDNA; 228 BP.
XX
AC AAV21723;
XX
DT 17-AUG-1998 (first entry)
XX
DE Humanised vector splice polyA signal sequence.
XX
KW Vector; vaccine; tumour; antigen; ds.
XX
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..17
FT /*tag= a
FT /note= "potential site for IRES"
FT intron 18..120
FT /*tag= b
FT /number= 3
FT exon 121..228
FT /*tag= c
FT /number= 4
FT polyA_signal 176..181
FT /*tag= d
XX
XX WO9806863-A1.
XX
XX
PD 19-FEB-1998.
XX

PF 14-AUG-1997; 97WO-US14306.
XX
PR 14-AUG-1996; 96US-0023931.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson EL, Nelson PJ;
XX
DR WPI; 1998-159552/14.
XX
PT Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
PS Example 1; Page 27; 125pp; English.
XX
CC This DNA sequence comprises a combined 3' splice sequence and
CC polyA tail sequence cloned into the Xho and HindIII sites of
CC pGEM-7zf. Novel humanised vectors of the invention (see AAV21724,
CC AAV21727 and AAV21732-34) comprise a human-derived promoter or mammalian
CC homologue which is functional in mammalian target tissue and cells
CC and an acceptance site which accepts cDNA products from RT-PCR
CC cloning. They may also include the splice-polyA signal
CC sequence. After processing of mRNA, intron 3 is excised, and
CC the sequence contains duplicate stop codons in 2 reading frames.
CC A 3rd reading frame stop codon can be obtained by mutagenising
CC the internal ribosomal entry site (IRES). The vectors are used to
CC express target antigens, especially tumour antigens. They are
CC non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences, and selectively elicit
CC immune responses to the target sequences, with little or no immune
CC response to the other vector components. The target antigens are
CC expressed as intracellular polypeptides and are appropriately
CC presented on antigen presenting cells.
XX
SQ Sequence 228 BP; 47 A; 63 C; 62 G; 56 T; 0 other;

Query Match 94.4%; Score 214.4; DB 19; Length 228;
Best Local Similarity 99.1%; Pred. No. 1.3e-58;
Matches 226; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGGAT-GGGAGACCTG 59
DB 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGGATGgggggagacctg 60
QY 60 TAGTCAGAGCCCCGGGCGAGCAGCCCAATGCCGCTCTCCCTGCAGGATGAGTAGT 119
DB 61 tagtcagagccccgggcgagcagagcccaatgccgctctctccctgcaggatgagtagt 120
QY 120 GAGTGCCTCTCTGCGCCCTGGGAAGTTGCCACTCCAGTGCCACCAGCCCTTGTCCTTAATAA 179
DB 121 gagtgcctctctgccccgggagtgccactccagtgccccagcctgtgcctaataa 180
QY 180 AATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227
DB 181 aattaagttgcatactttgtctgactaggtgtcctctataatattat 228

RESULT 3
AAV21732
ID AAV21732 standard; cDNA; 1547 BP.
XX
AC AAV21732;
XX
DT 17-AUG-1998 (first entry)
XX
DE Humanised vector pITL-A.
XX
KW Vector; vaccine; tumour; antigen; plasmid pITL-A; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.

```

XX Key Location/Qualifiers
PH CDS 472..680
FT /*tag= c
FT /*note= "SupF gene"
FT misc_feature 686..1292
FT /*tag= d
FT /*note= "ColE1 origin of replication"
FT promoter 1311..1547
FT /*tag= e
FT /*note= "RANTES promoter"
XX
XX W09806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Claim 14; Page 56-57; 125pp; English.
XX
XX Plasmid pITL-A comprises a base vector for novel humanised
XX polynucleotide vectors. The vector in Escherichia coli DH
XX 10-beta/P3 is deposited as ATCC 98401. Novel vectors comprise a
XX human-derived promoter or mammalian homologue which is functional
XX in mammalian target tissue and cells and a sequence acceptance site
XX (see AAV21735-36) which accepts cDNA products from RT-PCR cloning.
XX They also contain minimal non-human components, such as a replication
XX origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
XX are necessary for production of the vector. The novel vectors are
XX used to express target antigens, especially tumour antigens. They
XX are non-replicating in mammalian cells but are capable of extended
XX stable expression of target sequences generating an immune response
XX in immunised individuals. The vectors selectively elicit immune
XX responses to the target sequences with little or no immune response
XX to the other components of the vectors. The target antigens are
XX expressed as intracellular polypeptides or peptides and, as such,
XX are processed as self polypeptides or peptides and appropriately
XX presented on antigen presenting cells.
XX
XX Sequence 1547 BP; 344 A; 408 C; 425 G; 370 T; 0 other;

Query Match 94.4%; Score 214.4; DB 19; Length 1547;
Best Local Similarity 99.1%; Pred. No. 2.8e-58;
Matches 226; Conservative 1; Indels 1; Gaps 1;

QY 1 GCCTTAAGGGCCATATGCTGAGTGATGCTTACCCAGCGGGGAT-GGGGAGACCTG 59
DB 232 gctctaaggggccatatggtgagtgatgcttgacccagcgggatggggagacctg 291
QY 60 TAGTCAGAGCCCGCGGACAGCACAGGCCAATGCCGTCTTCCCTGCAGGATGAGTAGT 119
DB 292 tagtcaagaccccccggcagcagcagcagcagcagcagcagcagcagcagcagcag 351
QY 120 GAGTGCCTCTCTGCGCCCTGGAGTTGCCACTGCCAGTCCACCGCCTGTCTTAATAA 179
DB 352 gagtgcctctctctgcccctggagttggcactccagtcgccaccagcctgtcctataa 411
QY 180 AATTAAGTTGCATCATTTTGTCTGACTAGTGTCTCTATATATATAT 227
DB 412 aattaagtgcatacatttctgtctgactaggtgtcctctataattat 459

```

```

RESULT 4
AAV21733
ID AAV21733 standard; cDNA; 1807 BP.
XX
XX AAV21733;
XX
XX 17-AUG-1998 (first entry)
XX
XX Humanised vector pITL-1.
XX
XX Vector; vaccine; tumour; antigen; plasmid pITL-1; ds.
XX
XX Chimeric - Homo sapiens.
XX Chimeric - Escherichia coli.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 472..680
XX /*tag= c
XX /*note= "SupF gene"
XX misc_feature 686..1292
XX /*tag= d
XX /*note= "ColE1 origin of replication"
XX promoter 1605..1847
XX /*tag= e
XX /*note= "RANTES promoter"
XX
XX W09806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Claim 14; Page 56; 125pp; English.
XX
XX Plasmid pITL-1 comprises a base vector for novel humanised
XX polynucleotide vectors. The vector in Escherichia coli DH
XX 10-beta/P3 is deposited as ATCC 98400. Novel vectors comprise a
XX human-derived promoter or mammalian homologue which is functional
XX in mammalian target tissue and cells and a sequence acceptance site
XX (see AAV21735-36) which accepts cDNA products from RT-PCR cloning.
XX They also contain minimal non-human components, such as a replication
XX origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
XX are necessary for production of the vector. The novel vectors are
XX used to express target antigens, especially tumour antigens. They
XX are non-replicating in mammalian cells but are capable of extended
XX stable expression of target sequences generating an immune response
XX in immunised individuals. The vectors selectively elicit immune
XX responses to the target sequences with little or no immune response
XX to the other components of the vectors. The target antigens are
XX expressed as intracellular polypeptides or peptides and, as such,
XX are processed as self polypeptides or peptides and appropriately
XX presented on antigen presenting cells.
XX
XX Sequence 1807 BP; 419 A; 463 C; 492 G; 433 T; 0 other;

Query Match 94.4%; Score 214.4; DB 19; Length 1807;
Best Local Similarity 99.1%; Pred. No. 3e-58;
Matches 226; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 GCCTAAGGCCCATATGCTGAGTCGATCCCTTGACCCAGCGGGGAT-GGGAGACCTG 59
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 232 gccttaagggccatagtgtgagtgcccttgacccagcgaggatgggggagacctg 291
 QY 60 TAGTCAGAGCCCCCGGCGAGCACAGGCCAATGCCGCTCTTCCCTGCGAGGATGAGTAGT 119
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 292 tagtcagagcccccgccgagcacagggccaaagccgctctccctcgaggatgagtagt 351
 QY 120 GAGTGCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTCCCGCCAGCCAGCCTTCTCCTAATAA 179
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 352 gagtgcctctctgcccctggaagtggccactccagtgcccaccagcctgtctctaataa 411
 QY 180 AATTAAGTTGCATATTTGTCTGACTAGGTGCTCTATATATATAT 227
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 412 aattaagttgcatctttgtctgactagggtctctataattat 459
 RESULT 5
 AAV21734
 ID AAV21734 standard; cDNA; 2308 BP.
 XX AC AAV21734;
 XX 17-AUG-1998 (first entry)
 XX Humanised vector pITL-1 GFP.
 DE Vector; vaccine; tumour; antigen; plasmid pITL-1 GFP;
 KW green fluorescent protein; ds.
 KW Chimeric - Homo sapiens.
 XX Chimeric - Escherichia coli.
 OS Chimeric - Aequorea victoria.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT CDS 20..734
 FT /*tag= a
 FT /product= green fluorescent protein
 FT CDS 973..1181
 FT /*tag= C
 FT /note= "SupF gene"
 FT misc_feature 1191..1793
 FT /*tag= d
 FT /note= "ColE1 origin of replication"
 FT promoter 2063..2308
 FT /*tag= e
 FT /note= "RANTES promoter"
 XX WO9806863-A1.
 PN 19-FEB-1998.
 XX 14-AUG-1997; 97WO-US14306.
 XX 14-AUG-1996; 96US-0023931.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Nelson EL, Nelson PJ;
 PI WPI; 1998-159552/14.
 XX Humanised polynucleotide vectors - comprising human derived promoter
 XX and sequence acceptance site, used for the production of vaccines
 PT Example 11; Page 57-58; 125pp; English.
 PS Plasmid pITL-1 GFP comprises base vector pITL-1 (see AAV21733) and
 CC a humanised green fluorescent protein (GFP) reporter sequence
 CC (see AAV21725). Novel humanised vectors comprise a human-derived
 CC promoter or mammalian homologue which is functional in mammalian
 CC target tissue and cells and a sequence acceptance site which

CC accepts cDNA products from RT-PCR cloning. The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX Sequence 2308 BP; 571 A; 601 C; 604 G; 532 T; 0 other;
 SQ
 Query Match 94.4%; Score 214.4; DB 19; Length 2308;
 Best Local Similarity 99.1%; Pred. No. 3.3e-58;
 Matches 226; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 GCCTAAGGCCCATATGCTGAGTCGATCCCTTGACCCAGCGGGGAT-GGGAGACCTG 59
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 733 gccttaagggccatagtgtgagtgcccttgacccagcgaggatgggggagacctg 792
 QY 60 TAGTCAGAGCCCCCGGCGAGCACAGGCCAATGCCGCTCTTCCCTGCGAGGATGAGTAGT 119
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 793 tagtcagagcccccgccgagcacagggccaaagccgctctccctcgaggatgagtagt 852
 QY 120 GAGTGCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTCCCGCCAGCCAGCCTTCTCCTAATAA 179
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 853 gagtgcctctctgcccctggaagtggccactccagtgcccaccagcctgtctctaataa 912
 QY 180 AATTAAGTTGCATATTTGTCTGACTAGGTGCTCTATATATATAT 227
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 913 aattaagttgcatctttgtctgactagggtctctataattat 960
 RESULT 6
 AAV21724
 ID AAV21724 standard; cDNA; 1425 BP.
 XX AC AAV21724;
 XX 17-AUG-1998 (first entry)
 XX Humanised vector pITL.
 DE Vector; vaccine; tumour; antigen; plasmid pITL; ds.
 KW Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT misc_feature 1..221
 FT /*tag= a
 FT /note= "stuffer sequence"
 FT polyA_site 22..481
 FT /*tag= b
 FT /note= "combined splice and polyA sequences"
 FT CDS 495..701
 FT /*tag= C
 FT /note= "SupF gene"
 FT misc_feature 712..1164
 FT /*tag= d
 FT /note= "ColE1 origin of replication"
 FT promoter 1177..1425
 FT /*tag= e
 FT /note= "RANTES promoter"
 XX WO9806863-A1.
 PN 19-FEB-1998.
 XX 14-AUG-1997; 97WO-US14306.

```

XX PR 14-AUG-1996; 96US-0023931.
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA Nelson EL, Nelson PJ;
XX PI WPI; 1998-159552/14.
XX DR Humanised polynucleotide vectors - comprising human derived promoter
XX PT and sequence acceptance site, used for the production of vaccines
XX PS Claim 14; Page 29-30; 125pp; English.
XX CC Plasmid pITL comprises a base vector for novel humanised
XX CC polynucleotide vectors. Such vectors comprise a human-derived
XX CC promoter or mammalian homologue which is functional in mammalian
XX CC target tissue and cells and a sequence acceptance site (see
XX CC AAV21735-36), which accepts cDNA products from RT-PCR cloning. They
XX CC also contain minimal non-human components, such as a replication
XX CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
XX CC are necessary for production of the vector, as well as human-derived
XX CC splice and polyA sequences (see AAV21723). The novel vectors are
XX CC used to express target antigens, especially tumour antigens. They
XX CC are non-replicating in mammalian cells but are capable of extended
XX CC stable expression of target sequences generating an immune response
XX CC in immunised individuals. The vectors selectively elicit immune
XX CC responses to the target sequences with little or no immune response
XX CC to the other components of the vectors. The target antigens are
XX CC expressed as intracellular polypeptides or peptides and, as such,
XX CC are processed as self polypeptides or peptides and appropriately
XX CC presented on antigen presenting cells.
XX SQ Sequence 1425 BP; 325 A; 381 C; 404 G; 315 T; 0 Other;

Query Match 70.7%; Score 160.6; DB 19; Length 1425;
Best Local Similarity 84.2%; Pred. No. 3.1e-41;
Matches 218; Conservative 0; Mismatches 9; Indels 32; Gaps 2;

QY 1 GCCTTAAGGGCCATATGTTGAGTGATCCCTTGACCCAGCGGGGAT-GGGGAGACCTG 59
DB 223 gctttaaaggccatattggtgagtgatgcttgacccagcgggagtgaggacctg 282
QY 60 TAGTCAGAGCCCCCGGCGAGCAGACGCCAATGCCGTCCTCCCTCGCAG----- 109
DB 283 tagtcagagccccggcgagcagacagcccaatgcccgtctctccctgcaagtgtgtagtga 342
QY 110 -----GATGAGTAGTGAGTGAGTCCTCTCTGCGCCCTGGAAGTTGCC 148
DB 343 ctgcccgggtggatccctgtgacccctcccccagtgccctctcctgcccctggaagtggc 402
QY 149 ACTCAGTCCCGCCAGCAGCTTCTCTTAATAAAATTAAGTTGATCATTTTGTCTGACTAG 208
DB 403 actccagtgccacacagccttgctcctaataaaataaagttgcatcatttctgactag 462
QY 209 GTGTCCTCTATAATATAT 227
DB 463 gtgtcctctataattat 481

RESULT 7
AAV21727
ID AAV21727 standard; cDNA; 2125 BP.
XX AC AAV21727;
XX DT 17-AUG-1998 (first entry)
XX DE Humanised vector pITL-hHER/neu.
XX KW Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;
KW human; HER-2/neu; C-erbB-2; breast cancer; ds.

```

```

XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Escherichia coli.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX CDS 13..921
XX FT /*tag= a
XX FT /product= human HER-2/neu
XX FT polyA_site 922..1181
XX FT /*tag= b
XX FT /note= "combined splice and polyA sequences"
XX FT CDS 1195..1401
XX FT /*tag= c
XX FT /note= "SupF gene"
XX FT misc_feature 1412..1864
XX FT /*tag= d
XX FT /note= "COLE1 origin of replication"
XX FT promoter 1877..2125
XX FT /*tag= e
XX FT /note= "RANTES promoter"
XX XX WO9806863-A1.
XX PN 19-FEB-1998.
XX PD 14-AUG-1997; 97WO-US14306.
XX PF 14-AUG-1996; 96US-0023931.
XX PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA Nelson EL, Nelson PJ;
XX PI WPI; 1998-159552/14.
XX DR Humanised polynucleotide vectors - comprising human derived promoter
XX PT and sequence acceptance site, used for the production of vaccines
XX PS Example 7; Page 41-42; 125pp; English.
XX CC Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
XX CC a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
XX CC evaluate the toxicity of anti-tumour vaccination in rats, and in
XX CC phase I and phase II trials to evaluate polynucleotide vaccination
XX CC in advanced breast cancer. Novel humanised vectors, which can be
XX CC based on pITL, comprise a human-derived promoter or mammalian
XX CC homologue which is functional in mammalian target tissue and cells
XX CC and a sequence acceptance site which accepts cDNA products from
XX CC RT-PCR cloning. The vectors are non-replicating in mammalian cells
XX CC but are capable of extended stable expression of the target
XX CC sequence, generating an immune response in immunised individuals.
XX CC The vectors selectively elicit immune responses to the target
XX CC sequences with little or no immune response to the other components
XX CC of the vectors.
XX SQ Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 Other;

Query Match 70.7%; Score 160.6; DB 19; Length 2125;
Best Local Similarity 84.2%; Pred. No. 3.7e-41;
Matches 218; Conservative 0; Mismatches 9; Indels 32; Gaps 2;

QY 1 GCCTTAAGGGCCATATGTTGAGTGATCCCTTGACCCAGCGGGGAT-GGGGAGACCTG 59
DB 922 gctttaaaggccatattggtgagtgatgcttgacccagcgggagtgaggacctg 981
QY 60 TAGTCAGAGCCCCCGGCGAGCAGACGCCAATGCCGTCCTCCCTCGCAG----- 109
DB 982 tagtcagagccccggcgagcagacagcccaatgcccgtctctcctgcaagtgtgtagtga 1041
QY 110 -----GATGAGTAGTGAGTGAGTCCTCTCTGCGCCCTGGAAGTTGCC 148

```


FT fucosyltransferase to be inserted into
FT exon 1 of the hGH gene"

US5625124-A.

29-APR-1997.

11-JUL-1994; 94US-0273411.

11-JUL-1994; 94US-0273411.

(UNIW) UNIV WASHINGTON.

Falk P, Gordon JI;

WPI; 1997-258275/23.

Animal model for Helicobacter pylori infection - comprising
transgenic mouse expressing human enzyme promoting intestinal
adhesion

Example 2; Columns 25-28; 24pp; English.

A claimed transgenic mouse expresses, in its intestinal epithelial
cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
called alpha 1,3/4 FT). The enzyme is expressed under the
control of a gut epithelial cell-specific promoter and Helicobacter
pylori adheres to the transgenic cells. The transgenic mouse and
intestinal epithelial cells from it are useful as models for screening
compounds for the ability to inhibit adhesion of H. pylori to gut
epithelial cells. The first 617 nucleotides of the present chimeric
sequence encode the promoter from rat liver fatty acid binding protein
(fabp1) which can direct foreign gene expression to the pit cell
lineage of the mouse gastric epithelium, to proliferating and non-
proliferating cells in intestinal crypts, as well as to the four
principal differentiated cell lineages along the crypt-to-villus axis
of the small intestine. The remainder of the chimeric sequence
corresponds to exon 1 of the human growth hormone (hGH) gene, into
which a sequence coding for a human fucosyltransferase (hFT) can be
inserted. No hGH will be produced because the initiator Met codon
and the first translation stop codon will be from the hFT sequence
and there is no ribosome re-entry sequence. The hGH exon 1 ensures
efficient splicing of the transgene primary transcript, improves
stability of the cytoplasmic hFT mRNA and allows transgene
expression to be monitored by in situ hybridisation using a
digoxigenin-labelled hGH oligonucleotide.

Sequence 2771 BP; 666 A; 718 C; 704 G; 683 T; 0 other;

Query Match 46.4%; Score 105.4; DB 18; Length 2771;
Best Local Similarity 99.1%; Pred. No. 1.3e-23;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCTGGCCCTGGAGTTGCCACTCCAGTGCACCCAGCCCTGTCTTAATAAA 180

DB 2177 agtgcctctcctggccttggaagtggcaactccagtgccaccagcctgtgctctaataaa 2236

QY 181 ATTAAGTTGCATCTTTCTGACTAGGTGCTCTATAATATTAT 227

DB 2237 attaatgtgcatactttgtctgactaggtgctctctaataattat 2283

RESULT 10

AAC83268

ID AAC83268 standard; DNA; 4456 BP.

XX AAC83268;

AC AAC83268;

DT 16-MAR-2001 (first entry)

XX

DE Nucleotide sequence of HS2-beta globin-hGH transgene.

XX Erythropoietin-inducible transgene; human; sheep; juvenile beta globin;
KW human growth hormone; hGH; osteopathic; osteopaenia; osteoporosis;
KW hypersensitive element; HS; locus control region; ds.

OS Chimeric - Homo sapiens.

OS Chimeric - Ovis aries.

PN US6153427-A.

XX 28-NOV-2000.

XX 12-OCT-1994; 94US-0321686.

PR 12-OCT-1994; 94US-0321686.

XX (UYNE-) UNIV NORTHEASTERN OHIO.

XX Schneider GB, King D;

XX WPI; 2001-060092/07.

XX New DNA construct comprising an erythropoietin-inducible sequence, an
PT erythropoietin-specific enhancer sequence and coding sequence of interest
PT useful for treating osteopenia, particularly osteoporosis -

PS Claim 10; Fig 1B; 19pp; English.

XX This invention relates to a DNA construct comprising an

CC erythropoietin-inducible sequence represented by the present sequence. The
CC HS2-beta globin-hGH transgene comprises the hypersensitive site (HS2)
CC element from the human beta-globin locus control region, the juvenile
CC beta globin promoter isolated from sheep, and the human growth hormone
CC (hGH) gene. The transgene results in osteopathic activity, when used in
CC gene therapy. The DNA construct is useful for the treatment of
CC osteopaenias, such as osteoporosis.

XX Sequence 4456 BP; 1238 A; 978 C; 1076 G; 1164 T; 0 other;

Query Match 46.4%; Score 105.4; DB 22; Length 4456;
Best Local Similarity 99.1%; Pred. No. 1.6e-23;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCTGGCCCTGGAGTTGCCACTCCAGTGCACCCAGCCCTGTCTTAATAAA 180

DB 4324 agtgcctctcctggccttggaagtggcaactccagtgccaccagcctgtgctctaataaa 4383

QY 181 ATTAAGTTGCATCTTTCTGACTAGGTGCTCTATAATATTAT 227

DB 4384 attaatgtgcatactttgtctgactaggtgctctctaataattat 4430

RESULT 11

AAV33343

ID AAV33343 standard; DNA; 191 BP.

XX AAV33343;

XX 18-NOV-1998 (first entry)

DE Human growth hormone 3' UTR/poly(A) signal sequence.

XX Human interleukin 2; IL-2; CMV promoter; cationic lipid; DOTMA;
KW human growth hormone 3'-untranslated region; lipid cholesterol;
KW gene therapy; tumour; helper T cell; fever; fluid retention;
KW vascular leak syndrome; human growth hormone 3' UTR; ss.

OS Homo sapiens.

XX Key polyA_signal

XX Location/Qualifiers

FH 81..86

FT

FT XX /*Lag= a
 PN WO9834952-A2.
 XX 13-AUG-1998.
 PD 09-FEB-1998; 98WO-US02221.
 XX 23-JAN-1998; 98US-0012366.
 PR 10-FEB-1997; 97US-0039709.
 XX (GENE-) GENEMEDICINE INC.
 PA Bruno M, Muller S, Mumper R, Munger W, Ralston R;
 PI WPI; 1998-467159/40.
 XX Plasmid for expression of a human IL-2 coding sequence - and
 XX lipid/DNA delivery systems which have suitable positive/negative
 PT charge ratios and are useful in treating tumours
 PT
 XX Disclosure; Page 23; 77pp; English.
 FS
 XX The present sequence represents a human growth hormone 3' UTR/poly(A)
 CC signal sequence. The invention provides a pIL0697 plasmid comprising
 CC of a CMV promoter/enhancer transcriptionally linked to a codon
 CC optimised human IL-2 coding sequence (AAV3342) and the present human
 CC growth hormone 3'-untranslated region/poly(A) signal. Also, a 5' UTR
 CC (AAV3344) may be present between the 3' of the promoter region and the
 CC 5' of the IL-2 coding region. The invention also provides a composition
 CC for delivery and expression of a human IL-2 coding sequence in mammals,
 CC comprising: (a) a cationic lipid e.g. DOTMA, a neutral co-lipid e.g.
 CC lipid cholesterol and (b) a plasmid as described above. The
 CC plasmid/composition may be used in gene therapy, especially therapy of
 CC tumours, as IL-2 is involved in stimulating proliferation of helper T
 CC cells. The claimed antitumour effect of the above formulations does
 CC not only depend on the expression level of IL-2 but is also a function
 CC of the non-DNA formulation components, with the effect being greater
 CC than merely additive. Also, as administration of high doses of IL-2
 CC protein results in significant toxicity-related side effects, e.g.
 CC fever, fluid retention and vascular leak syndrome, the present
 CC composition is claimed to avoid such problems.
 XX
 SQ Sequence 191 BP; 39 A; 47 C; 56 G; 49 T; 0 other;

OS Homo sapiens.
 XX WO9947678-A2.
 PN 23-SEP-1999.
 XX 12-MAR-1999; 99WO-US05394.
 PF 19-MAR-1998; 98US-0078654.
 PR (GENE-) GENEMEDICINE INC.
 XX Nordstrom J, Pericle F, Rolland A, Ralston R;
 PI WPI; 1999-562116/47.
 XX New plasmids containing an interferon-alpha coding sequence, used for
 PT the treatment of a mammalian condition of disease, particularly cancer
 PT
 XX Disclosure; Page 27; 137pp; English.
 FS
 XX This sequence represents the 3' untranslated region from a human
 CC growth hormone gene. The invention relates to a novel plasmid comprising
 CC a cytomegalovirus (CMV) promoter transcriptionally linked with an
 CC interferon alpha (IFN-alpha) coding sequence, and a growth hormone
 CC 3'-untranslated region (UTR). The plasmids can be used for treating
 CC a mammalian condition of disease, particularly cancer.
 XX
 SQ Sequence 191 BP; 39 A; 47 C; 56 G; 49 T; 0 other;

Query Match 42.3%; Score 96; DB 20; Length 191;
 Best Local Similarity 99.1%; Pred. No. 4.5e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 121 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTGCCTGCTGTCTTAATAAA 180
 |||||||
 Db 27 agtgcctctctgcccgtggaattgccactccagtgcccaccagcctgtcctaataa 86
 |||||||

Oy 181 ATTAAGTTGCATCATTTTGTCTGCTAGTGTGCC-TCTATAATATAT 227
 |||||||
 Db 87 attaaagtgcacatctttgtctgactaggtgtccttctaataattat 134
 |||||||

RESULT 13
 AAZ50393
 ID AAZ50393 standard; DNA; 191 BP.
 XX
 AC AAZ50393;
 XX 18-MAY-2000 (first entry)
 DT Human growth hormone 3'UTR.
 DE
 XX Human growth hormone 3'UTR; anti-angiogenic agent; expression plasmid;
 XX cancer; gene expression; mRNA stability; transfection; tumour activity;
 KW cytotaxtic; gene therapy; lung metastatic tumour; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 XX 3'UTR 1..100
 FT /*tag= a
 FT polyA_signal 81..86
 FT /*tag= b
 XX
 XX WO200006759-A2.
 XX 10-FEB-2000.
 PD 20-JUL-1999; 99WO-US16388.
 XX

PR 27-JUL-1998; 98US-0094375.
 XX (VALE-) VALENTIS INC.
 XX
 PI Min W, Szymanski P, Mehrens D, Ralston R, Sullivan S;
 XX
 XX WPI; 2000-183133/16.
 XX
 PT Plasmids comprising tissue specific transcription elements linked to an
 PT anti-angiogenic gene is useful transfection of cells and treatment of,
 PT e.g. Cancer
 XX
 PS Disclosure; Page 28; 103pp; English.
 XX
 CC The present sequence is the 3'UTR of human growth hormone along with
 CC 3'flanking sequence. This sequence can be linked immediately following
 CC natural translation termination codon for a nucleotide encoding anti
 CC -angiogenic agent. This was used in the construction of expression
 CC plasmid incorporating an anti-angiogenic agent for the treatment of
 CC mammalian diseases, especially cancer. This 3'UTR influences gene
 CC expression by controlling the accuracy and efficiency of RNA processing,
 CC mRNA stability and translation. The plasmids can be used for (in vivo)
 CC transfection of a cell in situ in order to modulate tumour activity.
 CC Anti-angiogenic gene inhibits growth of solid tumour and lung metastatic
 CC tumours by intravenous or intramuscular delivery.
 XX
 SQ Sequence 191 BP; 39 A; 47 C; 56 G; 49 T; 0 other;

Query Match 42.3%; Score 96; DB 21; Length 191;
 Best Local Similarity 99.1%; Pred. No. 4.5e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 121 AGTGGCTCTCTGCGCCCTGGAAGTTGCCACTTCCAGTCCGCCACCGCCTTGCTTAATAAA 180
 Db 27 agtgcctctctgcccctggagttgccactccagtgccaccagcctgtctctaataaa 86
 Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTGCC-TCTATAATATTAT 227
 Db 87 attaatgtgcacattttgtctgactaggtgtcctctctataattat 134

RESULT 14
 AAF76860
 ID AAF76860 standard; cDNA; 563 BP.
 XX
 AC AAF76860;
 XX
 DT 14-MAY-2001 (first entry)
 XX
 DE Human secreted protein cDNA #18.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic;
 KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
 KW immunostimulant; cytostatic; cardiac; vascular; anti-angiogenic;
 KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
 KW antialzheimers; antiparkinsonian; antimicrobial; vulnery; gene therapy;
 KW immune disorder; hyperproliferative; cardiovascular; angiogenic;
 KW neurological; infection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200112776-A2.
 XX
 PD 22-FEB-2001.
 XX
 PF 15-AUG-2000; 2000WO-US22350.
 XX
 PR 16-AUG-1999; 99US-0148759.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;

XX WPI; 2001-244245/25.
 DR P-PSDB; AAB70079.
 XX
 PT Nucleic acids encoding 18 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 PS Claim 1; Page 358; 380pp; English.
 XX
 CC The present sequence is one of 18 nucleic acid molecules encoding novel
 CC human secreted proteins. The nucleic acids and proteins may be used in
 CC the prevention, diagnosis and treatment of diseases including immune
 CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and
 CC human immunodeficiency virus (HIV) infections), hyperproliferative
 CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration
 CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples. The polypeptides may also be
 CC used as antigens in the production of antibodies and in assays to
 CC identify modulators of protein expression and activity.
 XX
 SQ Sequence 563 BP; 127 A; 139 C; 176 G; 121 T; 0 other;

Query Match 42.3%; Score 96; DB 22; Length 563;
 Best Local Similarity 99.1%; Pred. No. 6.8e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 121 AGTGGCTCTCTGCGCCCTGGAAGTTGCCACTTCCAGTCCGCCACCGCCTTGCTTAATAAA 180
 Db 349 agtgcctctctgcccctggagttgccactccagtgccaccagcctgtctctaataaa 408
 Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTGCC-TCTATAATATTAT 227
 Db 409 attaatgtgcacattttgtctgactaggtgtcctctctataattat 456

RESULT 15
 AAX15886
 ID AAX15886 standard; DNA; 1441 BP.
 XX
 AC AAX15886;
 XX
 DT 13-MAY-1999 (first entry)
 XX
 DE NotI/SalI psk-GHRH plasmid fragment containing GHRH DNA sequence.
 XX
 KW Human; growth hormone releasing hormone; GHRH; osteoporosis; cachexia;
 KW growth disorder; burn; sepsis; trauma; atherogenic; cardiovascular;
 KW chronic obstructive pulmonary disease; atherosclerotic; cerebrovascular;
 KW peripheral vascular disease; haemophilia; muscular atrophy; aging;
 KW muscular dystrophy; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9905300-A2.
 XX
 PD 04-FEB-1999.
 XX
 PF 24-JUL-1998; 98WO-US15434.
 XX
 PR 20-OCT-1997; 97US-0062608.
 PR 24-JUL-1997; 97US-0053609.
 XX
 PA (GENE-) GENEMEDICINE INC.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 07:04:33 ; Search time 221.34 Seconds
(without alignments)
232.269 Million cell updates/sec

Title: US-09-242-202A-10
Perfect score: 227
Sequence: 1 GCCTTAAGGCCATATGGTG.....GGTGCTCTATAATATTAT 227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105.4	46.4	2771	1 US-08-273-411-5	Sequence 5, Appli
2	105.4	46.4	4456	3 US-08-321-686B-1	Sequence 1, Appli
3	96	42.3	191	3 US-09-012-366-4	Sequence 4, Appli
4	96	42.3	2086	3 US-08-589-028-9	Sequence 9, Appli
5	96	42.3	2086	3 US-08-784-582-9	Sequence 9, Appli
6	96	42.3	2086	4 US-08-785-271-9	Sequence 9, Appli
7	93.4	41.1	3653	4 US-08-973-334-1	Sequence 1, Appli
8	93.4	41.1	3653	4 US-09-563-869A-1	Sequence 1, Appli
9	93.4	41.1	3653	4 US-08-549-489-1	Sequence 1, Appli
10	72	31.7	792	3 US-08-887-674E-9	Sequence 1, Appli
11	60.6	26.7	1884	3 US-08-784-582-70	Sequence 70, Appl
12	60.6	26.7	2356	3 US-08-784-582-72	Sequence 72, Appl
c 13	42	18.5	42	3 US-08-068-754-3	Sequence 3, Appli
c 14	42	18.5	42	4 US-09-325-926-3	Sequence 3, Appli
c 15	42	18.5	51	2 US-08-453-024-10	Sequence 10, Appl
c 16	42	18.5	51	4 US-08-070-162-10	Sequence 10, Appl
c 17	39	17.2	46	4 US-09-173-043-10	Sequence 10, Appl
c 18	39	17.2	46	4 US-09-209-525-15	Sequence 15, Appl
c 19	39	17.2	153	4 US-09-209-525-16	Sequence 16, Appl
20	32.6	14.4	755	1 US-08-468-824-5	Sequence 5, Appli
c 21	31.8	14.0	3331	4 US-09-042-785A-1	Sequence 1, Appli
22	31.2	13.7	1758	3 US-09-191-171-6	Sequence 6, Appli
23	31.2	13.7	1758	4 US-09-385-707-6	Sequence 6, Appli
24	31.2	13.7	2220	2 US-08-864-224-1	Sequence 1, Appli
25	31.2	13.7	3853	3 US-08-801-092-5	Sequence 5, Appli
26	31.2	13.7	4026	3 US-08-801-092-19	Sequence 19, Appl
27	31.2	13.7	4249	3 US-08-801-092-33	Sequence 33, Appl

28	31.2	13.7	4283	1 US-08-343-401A-3	Sequence 3, Appli
29	31.2	13.7	4283	1 US-08-445-265A-1	Sequence 1, Appli
30	31.2	13.7	4283	3 US-08-990-442-1	Sequence 1, Appli
31	31.2	13.7	4326	4 US-08-760-615-7	Sequence 7, Appli
32	31.2	13.7	4328	4 US-09-132-808-1	Sequence 1, Appli
33	31.2	13.7	4328	4 US-08-910-647-2	Sequence 2, Appli
34	31.2	13.7	4328	4 US-08-910-647-2	Sequence 4, Appli
35	31.2	13.7	4928	1 US-08-345-913-1	Sequence 1, Appli
36	31.2	13.7	4928	3 US-08-818-562-1	Sequence 1, Appli
37	31.2	13.7	4965	2 US-08-564-313-1	Sequence 1, Appli
38	31.2	13.7	4965	5 PCT-US94-06089-1	Sequence 1, Appli
39	31.2	13.7	5107	4 US-08-910-647-3	Sequence 3, Appli
40	31.2	13.7	5653	1 US-08-073-836-3	Sequence 3, Appli
41	31.2	13.7	5653	1 US-08-235-277-1	Sequence 1, Appli
42	31.2	13.7	5676	2 US-08-663-998-3	Sequence 3, Appli
43	31.2	13.7	5682	2 US-08-663-998-4	Sequence 4, Appli
44	31.2	13.7	5703	1 US-08-467-420A-50	Sequence 50, Appl
45	31.2	13.7	5703	1 US-08-470-110A-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-08-273-411-5
; Sequence 5, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; CITY: Atlanta
; STREET: 1100 Peachtree Street, Suite 2800
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PUBLICATION INFORMATION:
; AUTHORS: Sweetser, et al.
; JOURNAL: Genes & Dev.
; VOLUME: 2
; PAGES: 1318-1332
; DATE: 1988
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 561 TO 629
; PUBLICATION INFORMATION:

```

; AUTHORS: Seeburg, et al.
; JOURNAL: DNA
; VOLUME: 1
; PAGES: 239-249
; DATE: 1982
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 620 TO 2771
; PUBLICATION INFORMATION:
; AUTHORS: Sweetser, et al.
; JOURNAL: J. Biol. Chem.
; VOLUME: 261
; PAGES: 5553-5561
; DATE: 1986
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 617
US-08-273-411-5

Query Match 46.4%; Score 105.4; DB 1; Length 2771;
Best Local Similarity 99.1%; Pred. No. 5.4e-24;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTGTGCTTAATAAA 180
Db 2177 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTGTGCTTAATAAA 2236

QY 181 ATTAAGTTGCATCATTTGCTGACTAGTGTGCTCTATAATATAT 227
Db 2237 ATTAAGTTGCATCATTTGCTGACTAGTGTGCTCTATAATATAT 2283

RESULT 2
US-08-321-686B-1
; Sequence 1, Application US/08321686B
; Patent No. 6153427
; GENERAL INFORMATION:
; APPLICANT: King, Donna
; APPLICANT: Schneider, Gary B.
; TITLE OF INVENTION: Treatment of
; TITLE OF INVENTION: Osteopenias and No. 6153427-Human Transgenic
; TITLE OF INVENTION: Mammals Expressing Therapeutic
; TITLE OF INVENTION: Transgenes
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wallenstein & Wagner, Ltd.
; STREET: 311 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch
; MEDIUM TYPE: 800 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: WordPerfect 5.1
; SOFTWARE: Simple Text ASCII (IBM format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,686B
; FILING DATE: October 12, 1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA: not
; PRIOR APPLICATION DATA: applicable
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry, Alan L.
; REGISTRATION NUMBER: 30,819
; REFERENCE/DOCKET NUMBER: 1017P021
; TELEPHONE: (312)554-3300
; TELEFAX: (312)554-3301
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4456 bp
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
US-08-321-686B-1

Query Match 46.4%; Score 105.4; DB 3; Length 4456;
Best Local Similarity 99.1%; Pred. No. 6.4e-24;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTGTGCTTAATAAA 180
Db 4324 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTGTGCTTAATAAA 4383

QY 181 ATTAAGTTGCATCATTTGCTGACTAGTGTGCTCTATAATATAT 227
Db 4384 ATTAAGTTGCATCATTTGCTGACTAGTGTGCTCTATAATATAT 4430

RESULT 3
US-09-012-366-4
; Sequence 4, Application US/09012366
; Patent No. 6034072
; GENERAL INFORMATION:
; APPLICANT: Robert Ralston
; APPLICANT: Susanne Muller
; APPLICANT: Russ Mumper
; APPLICANT: William Munger
; APPLICANT: Maria Bruno
; TITLE OF INVENTION: IL-2 GENE EXPRESSION AND
; TITLE OF INVENTION: DELIVERY SYSTEMS AND USES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,366
; FILING DATE: January 23, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/039,709
; FILING DATE: February 10, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkman, Charles S.
; REGISTRATION NUMBER: 38,077
; REFERENCE/DOCKET NUMBER: 230/214
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-012-366-4

Query Match 42.3%; Score 96; DB 3; Length 191;
Best Local Similarity 99.1%; Pred. No. 1.8e-21;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTGTGCTTAATAAA 180

```

Db 27 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCACCCAGCGCTTGCTCTAATAAA 86
 QY 181 ATTAAGTTCATCATTTTCTCTGACTAGGTGCC-TCTATAATATTAT 227
 Db 87 ATTAAGTTCATCATTTTCTCTGACTAGGTGCC-TCTATAATATTAT 134

RESULT 4
 US-08-589-028-9
 ; Sequence 9, Application US/08589028
 ; Patent No. 6087129
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6087129mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quaade, Christian
 ; APPLICANT: Kruse, Fred
 ; TITLE OF INVENTION: Recombinant Expression of Proteins From
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P. O. Box 4433
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 77210-4433
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/589,028
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Highlander, Steven L.
 ; REGISTRATION NUMBER: 47,642
 ; REFERENCE/DOCKET NUMBER: UTSD:426\HYL
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (512) 474-7577
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2086 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-589-028-9

Query Match 42.3%; Score 96; DB 3; Length 2086;
 Best Local Similarity 99.1%; Pred No. 4.4e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 121 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCACCCAGCGCTTGCTCTAATAAA 180
 Db 1559 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCACCCAGCGCTTGCTCTAATAAA 1618
 QY 181 ATTAAGTTCATCATTTTCTCTGACTAGGTGCC-TCTATAATATTAT 227
 Db 1619 ATTAAGTTCATCATTTTCTCTGACTAGGTGCC-TCTATAATATTAT 1666

RESULT 5
 US-08-784-582-9
 ; Sequence 9, Application US/08784582
 ; Patent No. 6110707
 ; GENERAL INFORMATION:

; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6110707mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quaade, Christian
 ; APPLICANT: Kruse, Fred
 ; APPLICANT: McGarry, Dennis
 ; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 ; NUMBER OF SEQUENCES: 79
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P. O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,582
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/028,427
 ; FILING DATE: 15-OCT-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/589,028
 ; FILING DATE: 19-JAN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Highlander, Steven L.
 ; REGISTRATION NUMBER: 37,642
 ; REFERENCE/DOCKET NUMBER: UTSD:514
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2086 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-784-582-9

Query Match 42.3%; Score 96; DB 3; Length 2086;
 Best Local Similarity 99.1%; Pred No. 4.4e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 121 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCACCCAGCGCTTGCTCTAATAAA 180
 Db 1559 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCACCCAGCGCTTGCTCTAATAAA 1618
 QY 181 ATTAAGTTCATCATTTTCTCTGACTAGGTGCC-TCTATAATATTAT 227
 Db 1619 ATTAAGTTCATCATTTTCTCTGACTAGGTGCC-TCTATAATATTAT 1666

RESULT 6
 US-08-785-271-9
 ; Sequence 9, Application US/08785271
 ; Patent No. 6194176
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6194176mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quaade, Christian

```

: APPLICANT: Kruse, Fred
:
: TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
:
: TITLE OF INVENTION: SECRETORY CELL LINES
:
: NUMBER OF SEQUENCES: 56
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Arnold, White & Durkee
:
: STREET: P.O. Box 4433
:
: CITY: Houston
:
: STATE: Texas
:
: COUNTRY: USA
:
: ZIP: 77210
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/785,271
:
: FILING DATE: Concurrently Herewith
:
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: US 08/589,028
:
: FILING DATE: 19-JAN-1996
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Highlander, Steven L.
:
: REGISTRATION NUMBER: 37,642
:
: REFERENCE/DOCKET NUMBER: UTSD:513
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 512/418-3000
:
: TELEFAX: 512/474-7577
:
: INFORMATION FOR SEQ ID NO: 9:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 2086 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: IS-08-785-271-9

```

	Query Match	Best Local Similarity	42.3%	Score 96;	DB 4;	Length 2086;
	Matches 107;	Conservative 0;	Mismatches 0;	Pred. No. 4.4e-21;		
					Indels 1;	Gaps 1;
QY	121	AGTGGCCTCTCTGGCCCTGGAAAGTTGCCACTCCAGTCCACGAGCCCTGTCTCTAATAAA	180			
Db	1559	AGTGGCCTCTCTGGCCCTGGAAAGTTGCCACTCCAGTCCACGAGCCCTGTCTCTAATAAA	1618			
QY	181	ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCC-TCATATAATATTAT	227			
Db	1619	ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTAATAATATTAT	1666			

```

7
RESULT
US-08-973-334-1
: Sequence 1, Application US/08973334
: Patent No. 6261551
: GENERAL INFORMATION:
: APPLICANT: Wilsson, James M.
: APPLICANT: Fisher, Krishna J.
: APPLICANT: Gao, Guang-Ping
: TITLE OF INVENTION: Recombinant Adenovirus and Adeno-
: TITLE OF INVENTION: Associated Virus, Cell Lines, and
: TITLE OF INVENTION: Methods of Production and Use
: TITLE OF INVENTION: Thereof
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Howson and Howson
: STREET: Box 457, 321 No. 6261551ristown Road
: CITY: Spring House
: STATE: PA
: COUNTRY: USA
: ZIP: 19477
: COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release 1.0 Version 1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,334
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,014
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/549,489
FILING DATE: 27-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN012CIPUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3653 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1521..2405
US-08-973-334-1

```

Query Match 41.1%; Score 93.4; DB 4; Length 3653;
Best Local Similarity 97.2%; Pred. No. 3.6e-20;
Matches 105; Conservative 0; Mismatches 2; Indels 1

QY	121	AGTGGCCTCTCCTGGCCCTGGAAGTTGGC	ACTCCACTGCGCCACGAGCCTTGCTTAATAAA	180
Db	2501	AGTGGCCTCTCCTGGCCCTGGAAGTTGGC	ACTCCACTGCGCCACGAGCCTTGCTTAATAAA	2560
QY	181	ATTAAGTTGCATCATTTTCTCTGACTAGGTGC	CTCTATATAATTAT	227
Db	2561	ATTAAGTTGCATCATTTTCTCTGACTAGGTGC	CTCTATATAATTAT	2608

RESULT 8
 US-09-563-869A-1
 ; Sequence 1, Application US/09563869A
 ; Patent No. 6270996
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Wilson, James M.
 ; Fisher, Krishna J.
 ; Gao, Guang-Ping
 ;
 ; TITLE OF INVENTION: Recombinant Adenovirus and Adeno-
 ; Associated Virus, Cell Lines, and
 ; Methods of Production and Use
 ; Thereof
 ;
 ; NUMBER OF SEQUENCES: 5
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Howson and Howson
 ; STREET: Box 457, 321 No. 6270996ristown Road
 ; CITY: Spring House
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19477
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release 1.0 Version 1.30
 ; CURRENT APPLICATION DATA:
 ;

; APPLICATION NUMBER: US/09/563,869A
; FILING DATE: 03-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/973,334
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/549,489
; FILING DATE: 27-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN012CIPUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1521..2405
; US-09-563-869A-1

Query Match 41.1%; Score 93.4; DB 4; Length 3653;
Best Local Similarity 97.2%; Pred. No. 3.6e-20;
Matches 105; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCACCGCCAGCCCTGTCTCTAATAAA 180
|||||
DB 2501 AGTGCCTCTCCTGGCCCTGGAAGTTGGCACTCCAGTGCACCGCCAGCCCTGTCTCTAATAAA 2560
|||||

QY 181 ATTAAGTTGCATCATTTTCTGTGACTAGGTGCC-TCTATAATATTAT 227
|||||
DB 2561 ATTAAGTTGNATCATTTTGTCTGACTAGGTGCCCTCTCTATAATATTAT 2608
|||||

RESULT 9
US-08-549-489-1
; Sequence 1, Application US/08549489
; Patent No. 6281010
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Gao, Guang-Ping
; TITLE OF INVENTION: No. 6281010el Adenovirus Gene Therapy Vehicle
; TITLE OF INVENTION: and Cell Line
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Box 457, 321 No. 6281010ristown Road
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,489
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/462,014
; FILING DATE: 08-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1521..2405
; US-08-549-489-1

Query Match 41.1%; Score 93.4; DB 4; Length 3653;
Best Local Similarity 97.2%; Pred. No. 3.6e-20;
Matches 105; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCACCGCCAGCCCTGTCTCTAATAAA 180
|||||
DB 2501 AGTGCCTCTCCTGGCCCTGGAAGTTGGCACTCCAGTGCACCGCCAGCCCTGTCTCTAATAAA 2560
|||||

QY 181 ATTAAGTTGCATCATTTTCTGTGACTAGGTGCC-TCTATAATATTAT 227
|||||
DB 2561 ATTAAGTTGNATCATTTTGTCTGACTAGGTGCCCTCTCTATAATATTAT 2608
|||||

RESULT 10
US-08-887-674E-9
; Sequence 9, Application US/08887674E
; Patent No. 6130092
; GENERAL INFORMATION:
; APPLICANT: Andre Lieber, Michael Strauss
; TITLE OF INVENTION: RIBOZYME GENE LIBRARY AND PRO-
; TITLE OF INVENTION: CESS FOR PREPARATION
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gabriel P. Katona
; STREET: 230 Park Avenue, Suite 2200
; CITY: New York
; STATE: NY
; COUNTRY: United States of America
; ZIP: 10169
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,674E
; FILING DATE: July 3, 1997
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: rRNA
; US-08-887-674E-9

Query Match 31.7%; Score 72; DB 3; Length 792;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCACCGCCAGCCCTGTCTCTAATAAA 180
|||||
DB 721 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCACCGCCAGCCCTGTCTCTAATAAA 780
|||||

QY 181 ATTAAGTTGCAT 192
 Db 781 ATTAAGTTGCAT 792

RESULT 11

US-08-784-582-70
 ; Sequence 70, Application US/08784582
 ; Patent No. 6110707
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6110707Mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quade, Christian
 ; APPLICANT: Kruse, Fred
 ; APPLICANT: McGarry, Dennis
 ; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 ; SECRETORY CELL LINES
 ; NUMBER OF SEQUENCES: 79
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,582
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/028,427
 ; FILING DATE: 15-OCT-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Highlander, Steven L.
 ; REGISTRATION NUMBER: 37,642
 ; REFERENCE/DOCKET NUMBER: UTSD:514
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 70:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1884 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

Query Match 26.7%; Score 60.6; DB 3; Length 1884;
 Best Local Similarity 84.2%; Pred. No. 5.9e-10;
 Matches 80; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 18 GTGAGTGGATCCCTTGACCCAGCGGGGAT-GGGGAGACCTGTAGTCAGAGCCCCCGGG 76
 Db 818 GTGAGTGGATCCCTTGACCCAGCGGGGATGGGGAGACCTGTAGTCAGAGCCCCCGGG 877

RESULT 12

US-08-784-582-72
 ; Sequence 72, Application US/08784582
 ; Patent No. 6110707
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6110707Mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quade, Christian
 ; APPLICANT: Kruse, Fred
 ; APPLICANT: McGarry, Dennis
 ; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 ; SECRETORY CELL LINES
 ; NUMBER OF SEQUENCES: 79
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,582
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/028,427
 ; FILING DATE: 15-OCT-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Highlander, Steven L.
 ; REGISTRATION NUMBER: 37,642
 ; REFERENCE/DOCKET NUMBER: UTSD:514
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 72:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2356 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

Query Match 26.7%; Score 60.6; DB 3; Length 2356;
 Best Local Similarity 84.2%; Pred. No. 6.4e-10;
 Matches 80; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 18 GTGAGTGGATCCCTTGACCCAGCGGGGAT-GGGGAGACCTGTAGTCAGAGCCCCCGGG 76
 Db 818 GTGAGTGGATCCCTTGACCCAGCGGGGATGGGGAGACCTGTAGTCAGAGCCCCCGGG 877

RESULT 13

US-08-068-754-3/c
 ; Sequence 3, Application US/08068754
 ; Patent No. 6023011
 ; GENERAL INFORMATION:
 ; APPLICANT: Vale, Wyle W.
 ; APPLICANT: Stenzel-Poore, Mary P.

APPLICANT: Koob, George F.
APPLICANT: Heinrichs, Stephen C.
TITLE OF INVENTION: CORTICOTROPIN-RELEASING FACTOR
TITLE OF INVENTION: OVERPRODUCING TRANSGENIC MICE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,754
FILING DATE: 28-MAY-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9397
TELEPHONE: 619-546-9392
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Oligonucleotide
US-08-068-754-3

Query Match 18.5%; Score 42; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 CCCTGGAAGTTGCCACTCCAGTCCCGCCAGCCAGCCTTGTCTCTAA 176
|||||
Db 42 CCCTGGAAGTTGCCACTCCAGTCCCGCCAGCCAGCCTTGTCTCTAA 1

RESULT 14
US-09-325-926-3/c
; Sequence 3, Application US/09325926
; Patent No. 6166287
; GENERAL INFORMATION:
; APPLICANT: Vale, Wyle W.
; Stenzel-Poore, Mary P.
; Koob, George F.
; Heinrichs, Stephen C.
; TITLE OF INVENTION: CORTICOTROPIN-RELEASING FACTOR
; OVERPRODUCING TRANSGENIC MICE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/325,926
FILING DATE: 04-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,754
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9397
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-9392
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Oligonucleotide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-325-926-3

Query Match 18.5%; Score 42; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 CCCTGGAAGTTGCCACTCCAGTCCCGCCAGCCAGCCTTGTCTCTAA 176
|||||
Db 42 CCCTGGAAGTTGCCACTCCAGTCCCGCCAGCCAGCCTTGTCTCTAA 1

RESULT 15
US-08-453-024-10/c
; Sequence 10, Application US/08453024
; Patent No. 5958707
; GENERAL INFORMATION:
; APPLICANT: de Vries, Jan
; APPLICANT: Jenh, Chung-Her
; APPLICANT: Narula, Satwant
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Human Interleukin-4 Antagonist/Agonist Screens
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,024
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,162
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: US07/869914
; FILING DATE: 16-APR-1992
; APPLICATION NUMBER: US07/770081
; FILING DATE: 03-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Dulak, No. 5958707man C.
; REGISTRATION NUMBER: 31,608
; REFERENCE/DOCKET NUMBER: JB0214K
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 201 822 7375
; TELEFAX: 201 822 7039
; TELEX: 219165
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-453-024-10

```

```

Query Match      18.5%  Score 42;  DB 2;  Length 51;
Best Local Similarity 100.0%;  Pred. No. 0.00011;
Matches 42;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy 165 GCCTTGCTCTAATAAATAAGTTGCATCATTTTGTCTGACT 206
      |||||||||||||||||||||||||||||||||||||||
Db 51 GCCTTGCTCTAATAAATAAGTTGCATCATTTTGTCTGACT 10

```

```

Search completed: January 17, 2002, 11:52:06
Job time: 17253 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2002, 11:48:13 ; Search time 10436.3 Seconds
(without alignments)
17,388 Million cell updates/sec

Title: US-09-242-202A-10_COPY_1_11
Perfect score: 11
Sequence: 1 GCCTTAAGGC 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Result No.	Score	Match	Length	DB	ID	Description
1	11	100.0	17	6	AX088916	Sequence
2	11	100.0	33	6	AR069952	Sequence
3	11	100.0	33	6	AX088922	Sequence
4	11	100.0	34	6	AR069934	Sequence
5	11	100.0	34	6	AR076077	Sequence
6	11	100.0	36	6	AR069933	Sequence
7	11	100.0	36	6	AR076076	Sequence
8	11	100.0	37	6	AR075081	Sequence
9	11	100.0	40	6	AR069956	Sequence
10	11	100.0	40	6	AR076090	Sequence
11	11	100.0	42	6	AR069950	Sequence
12	11	100.0	42	6	AR075083	Sequence
13	11	100.0	42	6	AR076083	Sequence
14	11	100.0	48	6	AR069955	Sequence
15	11	100.0	48	6	AR076089	Sequence
16	11	100.0	54	6	AR075075	Sequence
17	11	100.0	54	6	AX088904	Sequence
18	11	100.0	57	6	AR076086	Sequence
19	11	100.0	58	6	AR075077	Sequence
20	11	100.0	58	6	AR076085	Sequence
21	11	100.0	58	6	AX088906	Sequence
22	11	100.0	184	8	MSZ99636	Medicago sa
23	11	100.0	185	6	AX088919	Sequence
24	11	100.0	185	6	AX088920	Sequence
25	11	100.0	185	6	AX088921	Sequence
26	11	100.0	243	11	G05653	G05653 human STS W
27	11	100.0	268	11	G04304	G04304 human STS W
28	11	100.0	271	11	G64969	G64969 FBNI-25 Ran
29	11	100.0	274	1	AB000696	AB000696 Unidentif
30	11	100.0	275	1	AB000692	Unidentif
31	11	100.0	275	1	AB000693	Unidentif
32	11	100.0	275	1	AB000694	Unidentif
33	11	100.0	276	1	AB000695	Unidentif
34	11	100.0	349	11	G06533	G06533 human STS W
35	11	100.0	381	6	AX070615	Sequence
36	11	100.0	382	6	AX026147	Sequence
37	11	100.0	400	11	G14577	G14577 human STS S
38	11	100.0	407	11	HS222C3S	AL035374 H.sapiens
39	11	100.0	424	11	HSA290153	AJ290153 Homo sapi
40	11	100.0	425	11	HSA290145	AJ290145 Homo sapi
41	11	100.0	432	3	DROTNR82	K03275 D.melanogas
42	11	100.0	435	14	ESRVLTRL5	X95449 Endogenous
43	11	100.0	435	14	ESRVLTRL6	X95450 Endogenous
44	11	100.0	436	3	DROTNR81	M10908 D. melanoga
45	11	100.0	442	14	ENJRVLTR3	Z66533 Endogenous

ALIGNMENTS

RESULT 1	AX088916	AX088916	17 bp	DNA	PAT	17-MAR-2001
LOCUS	Sequence	Sequence	17	from Patent	WO0114534.	
DEFINITION	AX088916	AX088916				
ACCESSION	AX088916.1	GI:13397675				
VERSION						
KEYWORDS						
SOURCE	Bacillus subtilis.					
ORGANISM	Bacillus subtilis					

Bacteria: Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
1 (bases 1 to 17)
Thomas, M.D. and Brown, K.M.
Polypeptides having pectin acetyltransferase activity and nucleic
acids encoding same
Patent: WO 0114534-A 17 01-MAR-2001;
Novozymes Biotech, Inc. (US)
Location/Qualifiers
1. .17
/organism="Bacillus subtilis"
source

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
BASE COUNT      3 a      5 c      6 g      3 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  2 GCCTTAAGGCG 12

RESULT 4
AR069934      34 bp      DNA      PAT      18-FEB-2000
DEFINITION   Sequence 19 from patent US 5891701.
ACCESSION    AR069934
VERSION      AR069934.1 GI:7220822
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 34)
AUTHORS     Sloma,A. and Christianson,L.
TITLE       Nucleic acids encoding a polypeptide having protease activity.
JOURNAL     Patent: US 5891701-A 19 06-APR-1999;
FEATURES    Location/Qualifiers
            source
            1..34
            /organism="unknown"
BASE COUNT   5 a      10 c      11 g      8 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  3 GCCTTAAGGCG 13

RESULT 5
AR076077      34 bp      DNA      PAT      30-AUG-2000
DEFINITION   Sequence 19 from patent US 5958728.
ACCESSION    AR076077
VERSION      AR076077.1 GI:10002823
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 34)
AUTHORS     Sloma,A., Sternberg,D., Adams,L.F. and Brown,S.
TITLE       Methods for producing polypeptides in mutants of bacillus cells
JOURNAL     Patent: US 5958728-A 19 28-SEP-1999;
FEATURES    Location/Qualifiers
            source
            1..34
            /organism="unknown"
BASE COUNT   5 a      10 c      11 g      8 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  3 GCCTTAAGGCG 13

RESULT 6
AR069933/c
LOCUS        AR069933      36 bp      DNA      PAT      18-FEB-2000
DEFINITION   Sequence 18 from patent US 5891701.
ACCESSION    AR069933
VERSION      AR069933.1 GI:7220821
KEYWORDS
SOURCE       Unknown.
```

```
BASE COUNT      3 a      5 c      6 g      3 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  2 GCCTTAAGGCG 12

RESULT 2
AR069952      33 bp      DNA      PAT      18-FEB-2000
DEFINITION   Sequence 39 from patent US 5891701.
ACCESSION    AR069952
VERSION      AR069952.1 GI:7220840
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 33)
AUTHORS     Sloma,A. and Christianson,L.
TITLE       Nucleic acids encoding a polypeptide having protease activity
JOURNAL     Patent: US 5891701-A 39 06-APR-1999;
FEATURES    Location/Qualifiers
            source
            1..33
            /organism="unknown"
BASE COUNT   4 a      11 c      9 g      9 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  5 GCCTTAAGGCG 15

RESULT 3
AX088922      33 bp      DNA      PAT      17-MAR-2001
DEFINITION   Sequence 23 from Patent WO0114534.
ACCESSION    AX088922
VERSION      AX088922.1 GI:13397681
KEYWORDS
SOURCE       Bacillus subtilis.
ORGANISM     Bacillus subtilis
REFERENCE    1 (bases 1 to 33)
AUTHORS     Thomas,M.D. and Brown,K.M.
TITLE       Polypeptides having pectin acetylesterase activity and nucleic
JOURNAL     Patent: WO 0114534-A 23 01-MAR-2001;
FEATURES    Location/Qualifiers
            source
            1..33
            /organism="Bacillus subtilis"
            /db_xref="taxon:1423"
BASE COUNT   4 a      11 c      10 g      8 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 36)
AUTHORS Sloma,A. and Christianson,L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 18 06-APR-1999;
FEATURES Location/Qualifiers
source
BASE COUNT 11 a 12 c 10 g 3 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 22 GCCTTAAGGCG 12

RESULT 7
LOCUS AR076076 36 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 18 from patent US 5958728.
ACCESSION AR076076
VERSION AR076076.1 GI:10002822
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 36)
AUTHORS Sloma,A., Sternberg,D., Adams,L.F. and Brown,S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 18 28-SEP-1999;
FEATURES Location/Qualifiers
source
BASE COUNT 11 a 12 c 10 g 3 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 22 GCCTTAAGGCG 12

RESULT 8
LOCUS AR075081 37 bp DNA PAT 28-AUG-2000
DEFINITION Sequence 7 from patent US 5955310.
ACCESSION AR075081
VERSION AR075081.1 GI:10001833
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 37)
AUTHORS Widner,W., Sloma,A. and Thomas,M.D.
TITLE Methods for producing a polypeptide in a bacillus cell
JOURNAL Patent: US 5955310-A 7 21-SEP-1999;
FEATURES Location/Qualifiers
source
BASE COUNT 4 a 12 c 10 g 11 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 37;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 5 GCCTTAAGGCG 15

RESULT 9
LOCUS AR069956/c 40 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 46 from patent US 5891701.
ACCESSION AR069956
VERSION AR069956.1 GI:7220844
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Sloma,A. and Christianson,L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 46 06-APR-1999;
FEATURES Location/Qualifiers
source
BASE COUNT 7 a 14 c 13 g 6 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 40;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 38 GCCTTAAGGCG 28

RESULT 10
LOCUS AR076090/c 40 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 32 from patent US 5958728.
ACCESSION AR076090
VERSION AR076090.1 GI:10002836
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Sloma,A., Sternberg,D., Adams,L.F. and Brown,S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 32 28-SEP-1999;
FEATURES Location/Qualifiers
source
BASE COUNT 7 a 14 c 13 g 6 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 40;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 38 GCCTTAAGGCG 28

RESULT 11
LOCUS AR069950 42 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 37 from patent US 5891701.

ACCESSION AR069950
VERSION AR069950.1 GI:7220838
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 42)
AUTHORS Sloma, A. and Christianson, L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 37 06-APR-1999;
FEATURES Location/Qualifiers
source
1. 42
BASE COUNT 12 a 6 c 12 g 12 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 5 GCCTTAAGGC 15

RESULT 12
LOCUS AR075083 42 bp DNA PAT 28-AUG-2000
DEFINITION Sequence 9 from patent US 5955310.
ACCESSION AR075083
VERSION AR075083.1 GI:10001835
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 42)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Methods for producing a polypeptide in a bacillus cell
JOURNAL Patent: US 5955310-A 9 21-SEP-1999;
FEATURES Location/Qualifiers
source
1. 42
BASE COUNT 12 a 6 c 12 g 12 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 5 GCCTTAAGGC 15

RESULT 13
LOCUS AR076083 42 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 25 from patent US 5958728.
ACCESSION AR076083
VERSION AR076083.1 GI:10002829
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 42)
AUTHORS Sloma, A., Sternberg, D., Adams, L.F. and Brown, S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 25 28-SEP-1999;
FEATURES Location/Qualifiers
source
1. 42
BASE COUNT 12 a 6 c 12 g 12 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 5 GCCTTAAGGC 15

RESULT 14
LOCUS AR069955 48 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 45 from patent US 5891701.
ACCESSION AR069955
VERSION AR069955.1 GI:7220843
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 48)
AUTHORS Sloma, A. and Christianson, L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 45 06-APR-1999;
FEATURES Location/Qualifiers
source
1. 48
BASE COUNT 8 a 15 c 16 g 9 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 7 GCCTTAAGGC 17

RESULT 15
LOCUS AR076089 48 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 31 from patent US 5958728.
ACCESSION AR076089
VERSION AR076089.1 GI:10002835
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 48)
AUTHORS Sloma, A., Sternberg, D., Adams, L.F. and Brown, S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 31 28-SEP-1999;
FEATURES Location/Qualifiers
source
1. 48
BASE COUNT 8 a 15 c 16 g 9 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 7 GCCTTAAGGC 17

RESULT 16
LOCUS AR076089 48 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 31 from patent US 5958728.
ACCESSION AR076089
VERSION AR076089.1 GI:10002835
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 48)
AUTHORS Sloma, A., Sternberg, D., Adams, L.F. and Brown, S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 31 28-SEP-1999;
FEATURES Location/Qualifiers
source
1. 48
BASE COUNT 8 a 15 c 16 g 9 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 7 GCCTTAAGGC 17

Fri Jan 18 08:27:34 2002

us-09-242-202a-10_copy_1_11.rge

Page 5

Search completed: January 17, 2002, 11:48:13
Job time: 17105 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:02:18 ; Search time 599.86 Seconds
(without alignments)
15.721 Million cell updates/sec

Title: US-09-242-202a-10_COPY_1_11

Perfect score: 11

Sequence: 1 GCCTTAGGGC 11

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101:*

1:	/SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	11	100.0	11	AAV21736	Humanised vector 3
2	11	100.0	17	AAZ23342	Bacillus amyloliqu
3	11	100.0	17	AAD09919	PCR primer #1 for
4	11	100.0	17	AAF62642	amyQ promoter olig
5	11	100.0	33	AAZ23344	Bacillus amyloliqu
6	11	100.0	33	AAZ23344	Bacillus amyloliqu
7	11	100.0	33	AAZ23348	Bacillus amyL prom
8	11	100.0	33	AAV82419	Short consensus am
9	11	100.0	33	AAD09928	Short consensus am
10	11	100.0	33	AAF62648	amyQ promoter olig
11	11	100.0	34	AAV38595	PCR primer 19 used

c	12	11	100.0	34	20	AAV82401	Bacillus subtilis
c	13	11	100.0	36	19	AAV38594	PCR primer 18 used
	14	11	100.0	36	20	AAV82400	Bacillus subtilis
	15	11	100.0	37	20	AAZ23328	Bacillus lichenifo
	16	11	100.0	37	22	AAD09904	Bacillus lichenifo
	17	11	100.0	42	19	AAV38601	PCR primer 25 used
	18	11	100.0	42	20	AAZ23330	Bacillus amyloliqu
	19	11	100.0	42	20	AAV82417	Bacillus amyQ prom
	20	11	100.0	42	22	AAD09906	Bacillus amyloliqu
c	21	11	100.0	51	22	AAD09899	Synthetic polylink
	22	11	100.0	54	22	AAD09898	Synthetic polylink
	23	11	100.0	54	22	AAF62630	Polylinker #1. Sy
	24	11	100.0	58	22	AAD09900	Synthetic polylink
c	25	11	100.0	58	22	AAD09901	Synthetic polylink
	26	11	100.0	58	22	AAF62632	Polylinker #3. Sy
	27	11	100.0	63	20	AAZ34335	HIV polyA(2) compl
c	28	11	100.0	65	20	AAZ34334	HIV polyA(1) prime
	29	11	100.0	185	20	AAZ23322	B. thuringiensis w
	30	11	100.0	185	20	AAZ23323	B. amyloliquefacie
	31	11	100.0	185	20	AAZ23324	B. amyloliquefacie
	32	11	100.0	185	20	AAZ23325	Bacillus sp. amyQ
	33	11	100.0	185	22	AAD09922	Bacillus amyloliqu
	34	11	100.0	185	22	AAD09923	Bacillus amyloliqu
	35	11	100.0	185	22	AAD09924	Bacillus amyloliqu
	36	11	100.0	185	22	AAD09931	Bacillus amyloliqu
	37	11	100.0	185	22	AAF62645	Wild type amyQ pro
	38	11	100.0	185	22	AAF62646	Consensus amyQ pro
	39	11	100.0	185	22	AAF62647	Consensus amyQ pro
	40	11	100.0	227	19	AAV21719	Combined 3' splice
	41	11	100.0	228	19	AAV21723	Humanised vector s
	42	11	100.0	252	19	AAV21720	Humanised polynucl
	43	11	100.0	300	20	AAZ13626	Human gene express
c	44	11	100.0	308	21	AAZ32016	Plant microsatelli
c	45	11	100.0	338	21	AAZ31512	Plant microsatelli

ALIGNMENTS

RESULT 1

AAV21736

ID AAV21736 standard; cDNA; 11 BP.

XX AC AAV21736;

XX DT 17-AUG-1998 (first entry)

XX DE Humanised vector 3' sequence acceptance site.

XX KW Vector; vaccine; tumour; antigen; plasmid pITL; ss.

XX OS Synthetic.

XX PN WO9806863-A1.

XX PD 19-FEB-1998.

XX PF 14-AUG-1997; 97WO-US14306.

XX PR 14-AUG-1996; 96US-0023931.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Nelson EL, Nelson PJ;

XX DR WPI; 1998-159552/14.

XX PT Humanised polynucleotide vectors - comprising human derived promoter and sequence acceptance site, used for the production of vaccines

XX PS Claim 13; Page 104; 125pp; English.

XX CC This DNA sequence comprises a 5' acceptance site of novel humanised

CC polynucleotide vectors such as pITL (see AAV21724). Such vectors
CC comprise a human-derived promoter or mammalian homologue and a
CC sequence acceptance site that is designed to directionally accept
CC sequence specific products from RT-PCR based cloning strategies
CC via a unique site within an interrupted palindrome recognition
CC sequence for a restriction endonuclease which is incorporated into
CC the PCR primer. In this embodiment, the palindrome recognition
CC sequence is for BglI. The 5' acceptance site is given in AAV21735.
CC The novel vectors are used to express target antigens, especially
CC tumour antigens. They are non-replicating in mammalian cells but
CC are capable of extended stable expression of target sequences, and
CC generate immune responses to the target sequences with little or no
CC immune response to the other vector components.

XX SQ Sequence 11 BP; 2 A; 3 C; 4 G; 2 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 11;
Best Local Similarity 100.0%; Pred. No. 3e+02; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 GCCTTAAGGCG 11
Db 1 gccttaaggcg 11
|||||

RESULT 2
AAZ23342
ID AAZ23342 standard; DNA; 17 BP.
XX
AC AAZ23342;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus amyloliquefaciens consensus amyQ gene promoter PCR primer 2.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone;
KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;
KW alpha-amylase; ss.
XX
OS Synthetic.
OS Bacillus amyloliquefaciens.
PN WO9943835-A2.
XX
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
PR 26-FEB-1999; 98US-0031442.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 1999-561370/47.
XX
PT Production of polypeptide in Bacillus using specific promoters,
PT particularly for producing enzymes
XX
PS Example 18; Page 35; 90pp; English.

CC This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters.
CC The method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a PCR primer used to amplify
CC the Bacillus amyloliquefaciens consensus amyQ promoter.

XX SQ Sequence 17 BP; 3 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 3
AAD09919
ID AAD09919 standard; DNA; 17 BP.
XX
AC AAD09919;
XX
DT 12-SEP-2001 (first entry)
XX
DE PCR primer #1 for constructing mutated consensus amyQ promoter.
XX
KW Bacillus cell; tandem promoter; consensus promoter;
KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; PCR primer; ss.
XX
OS Bacillus amyloliquefaciens.
XX
PN US6255076-B1.
XX
PD 03-JUL-2001.
XX
PF 26-FEB-1999; 99US-0258377.
XX
PR 26-FEB-1999; 98US-0031442.
XX
PA (NOVO-) NOVOZYMES BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 2001-440518/47.
XX
PT Producing a polypeptide in a Bacillus strain comprises cultivating a
PT Bacillus cell, which contains a nucleic acid construct comprising a
PT tandem promoter and consensus promoters
XX
PS Example 18; Column 25; 54pp; English.

CC The present invention relates to a method for producing a polypeptide,
CC comprises cultivating a Bacillus cell, which contains a nucleic acid
CC construct comprising a tandem promoter and consensus promoters. The
CC Bacillus is cultivated in a medium conducive for the production of the
CC polypeptide. Each promoter sequence of the tandem promoter is operably
CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter
CC or subtilisin Carlsberg gene promoter have the sequence TTGACA for
CC the -35 region and TATAAT for the -10 region. The method further
CC comprises isolating the polypeptide from the cultivation medium. The
CC method is useful for producing a polypeptide in a Bacillus strain.
CC The present sequence is a PCR primer used in the construction of a
CC consensus mutated Bacillus amyloliquefaciens alpha-amylase (amyQ)
CC promoter. The primer generates a double-stranded fragment comprising
CC 137 bp of the upstream region of the amyQ promoter, used in the
CC exemplification of the invention.

XX SQ Sequence 17 BP; 3 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;

```
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 4
AAF62642
ID AAF62642 standard; DNA; 17 BP.
XX
AC AAF62642;
XX
DT 03-MAY-2001 (first entry)
XX
DE amyQ promoter oligonucleotide #3.
XX
KW Pectin acetyltransferase; degrade; plant cell wall; ss.
XX
OS Synthetic.
XX
PN US6184028-B1.
XX
PD 06-FEB-2001.
XX
PF 26-AUG-1999; 99US-0384305.
XX
PR 26-AUG-1999; 99US-0384305.
XX
PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX
PI Thomas MD, Brown KM;
XX
XX WPI; 2001-190946/19.
XX
Novel isolated polypeptide having pectin acetyltransferase activity useful
PT for degrading pectic substances and in degradation or modification of
PT acetylated pectins and plant cell walls
XX
XX Example 9; Column 28; 35pp; English.
XX
CC The present invention relates to Bacillus subtilis pectin
CC acetyltransferase protein. The invention is useful for degrading
CC a pectic substance. It is also useful for degrading soluble and
CC insoluble pectins with varying degrees of esterification,
CC clarification etc. The protein may be used alone or in
CC combination with other enzymes for the degradation or modification
CC of acetylated pectins, degradation or modification of plant cell walls.
XX
SQ Sequence 17 BP; 3 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 5
AA223344
ID AA223344 standard; DNA; 33 BP.
XX
AC AA223344;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus amyloliquefaciens short consensus amyQ promoter PCR primer 1.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone;
KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;
XX
```

```
KW alpha-amylase; SAVINASE; ss.
XX
OS Synthetic.
OS Bacillus amyloliquefaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
PR 26-FEB-1998; 98US-0031442.
XX
PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
XX WPI; 1999-561370/47.
DR
XX Production of polypeptide in Bacillus using specific promoters,
XX particularly for producing enzymes
PT
XX Example 19; Page 36; 90pp; English.
PS
XX This invention describes a novel method for the production of a
XX polypeptide in Bacillus using specific tandem or consensus promoters.
XX The method is used to produce homologous or particularly heterologous
XX proteins, particularly enzymes (specifically serine protease, maltogenic
XX alpha-amylase and pullulanase), but also hormones, antibodies, reporters
XX etc. The specified promoters provide increased expression of the sequence
XX which encodes the polypeptide of the invention. After incorporation of
XX the nucleic acid construct of the invention, any marker gene may be
XX deleted, resulting in a cell that is preferred for environmental and
XX regulatory regions. This sequence represents a PCR primer used to amplify
XX the Bacillus amyloliquefaciens short consensus amyQ promoter which is
XX used in the construction of a short consensus amyQ promoter-SAVINASE(RTM)
XX gene expression cassette.
XX
SQ Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 6
AA223348
ID AA223348 standard; DNA; 33 BP.
XX
AC AA223348;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus amyloliquefaciens short consensus amyQ promoter PCR primer 5.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone;
KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;
KW alpha-amylase; SAVINASE; ss.
XX
OS Synthetic.
OS Bacillus amyloliquefaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
```

PR 26-FEB-1998; 98US-0031442.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 FA Widner W, Sloma A, Thomas MD;
 XX WPI; 1999-561370/47.
 XX Production of polypeptide in Bacillus using specific promoters,
 PT particularly for producing enzymes -
 XX Example 25; Page 41; 90pp; English.
 XX This invention describes a novel method for the production of a
 CC polypeptide in Bacillus using specific tandem or consensus promoters.
 CC The method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a PCR primer used to amplify
 CC the Bacillus amyloliquefaciens short consensus amyQ trimer promoter which
 CC is used in the construction of a short consensus amyQ dimer promoter-
 CC SAVINASE(RTM) gene expression cassette.
 XX Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;
 SQ

Query Match 100.0%; Score 11; DB 20; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTTAAGGCG 11
 DB 2 gccttaaggcg 12
 |||||

RESULT 7
 ID AAV82419 standard; DNA; 33 BP.
 XX AAV82419;
 AC AAV82419;
 XX 12-APR-1999 (first entry)
 DT Bacillus amyL promoter PCR primer term1sfi.
 DE Protease; detergent; surfactant; leather processing; debittering;
 KW flavour; PCR; primer; amyL; amylase; promoter; ss.
 OS Synthetic.
 XX Bacillus amyloliquefaciens.
 XX WO9856927-A2.
 PN 17-DEC-1998.
 PD 09-JUN-1998; 98WO-US12005.
 PF 12-JUN-1997; 97US-0873479.
 PR (NOVO) NOVO NORDISK BIOTECH INC.
 XX Christianson L, Sloma A;
 PI WPI; 1999-080908/07.
 XX Novel protease from Bacillus subtilis LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing
 XX Example 12; Page 36; 77pp; English.

XX Primers term1sfi and 2isfi (see AAV82419-20) were used for the PCR
 CC amplification of the promoter region of the TERMAMYL amylase
 CC amyL gene of Bacillus amyloliquefaciens. Primer term1sfi was
 CC used to add an SfiI site to the 5' end of the promoter. The
 CC amplified promoter was inserted upstream of a recombinant
 CC Bacillus JP170 protease gene (see AAV82382) in vector pS2882-MCS,
 CC giving plasmid p170PERM. The gene encodes a protease (see AAV89547)
 CC useful in detergent compositions and in leather processing. Host
 CC cells, vectors and methods for recombinant production of the
 CC protease are provided.
 XX Sequence 33 BP; 4 A; 11 C; 9 G; 9 T; 0 other;
 SQ

Query Match 100.0%; Score 11; DB 20; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTTAAGGCG 11
 DB 5 gccttaaggcg 15
 |||||

RESULT 8
 ID AAD09925 standard; DNA; 33 BP.
 XX AAD09925;
 AC AAD09925;
 XX 12-SEP-2001 (first entry)
 DT Short consensus amyQ promoter amplifying PCR primer.
 DE Bacillus cell; tandem promoter; consensus promoter;
 KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
 KW cryIIIA promoter; subtilisin Carlsberg gene promoter; PCR primer; ss.
 XX Bacillus amyloliquefaciens.
 XX US6255076-B1.
 PN 03-JUL-2001.
 PD 26-FEB-1999; 99US-0258377.
 PF 26-FEB-1998; 98US-0031442.
 PR (NOVO-) NOVOZYMES BIOFTECH INC.
 PA Widner W, Sloma A, Thomas MD;
 XX WPI; 2001-440518/47.
 XX Producing a polypeptide in a Bacillus strain comprises cultivating a
 PT Bacillus cell, which contains a nucleic acid construct comprising a
 PT tandem promoter and consensus promoters -
 XX Example 19; Column 26; 54pp; English.
 XX The present invention relates to a method for producing a polypeptide,
 CC comprising cultivating a Bacillus cell, which contains a nucleic acid
 CC construct comprising a tandem promoter and consensus promoters. The
 CC Bacillus is cultivated in a medium conducive for the production of the
 CC polypeptide. Each promoter sequence of the tandem promoter is operably
 CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
 CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter
 CC or subtilisin Carlsberg gene promoter have the sequence TTGACA for
 CC the -35 region and TATAAT for the -10 region. The method further
 CC comprises isolating the polypeptide from the cultivation medium. The
 CC method is useful for producing a polypeptide in a Bacillus strain.

CC The present sequence is a PCR primer used in the construction of a
CC short consensus mutated Bacillus amyloliquefaciens alpha-amylase (amyQ)
CC promoter gene expression cassette, used in the exemplification of
CC the invention.
XX
SQ Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 9

AA009928
ID AAD09928 standard; DNA; 33 BP.

XX AC AAD09928;

DT 12-SEP-2001 (first entry)

XX Short consensus amyQ dimer promoter amplifying PCR primer.

XX Bacillus cell; tandem promoter; consensus promoter;

KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; PCR primer; ss.

XX OS Bacillus amyloliquefaciens.

XX PN US6255076-B1.

XX PD 03-JUL-2001.

XX PF 26-FEB-1999; 99US-0258377.

XX PR 26-FEB-1998; 98US-0031442.

XX PA (NOVO-) NOVOZYMES BIOTECH INC.

XX PI Widner W, Sloma A, Thomas MD;

XX DR WPI; 2001-440518/47.

XX Producing a polypeptide in a Bacillus strain comprises cultivating a
PT Bacillus cell, which contains a nucleic acid construct comprising a
PT tandem promoter and consensus promoters -

XX Example 25; Column 30; 54pp; English.

CC The present invention relates to a method for producing a polypeptide,
CC comprises cultivating a Bacillus cell, which contains a nucleic acid
CC construct comprising a tandem promoter and consensus promoters. The
CC Bacillus is cultivated in a medium conducive for the production of the
CC polypeptide. Each promoter sequence of the tandem promoter is operably
CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter
CC or subtilisin Carlsberg gene promoter have the sequence TTGACA for
CC the -35 region and TATAAT for the -10 region. The method further
CC comprises isolating the polypeptide from the cultivation medium. The
CC method is useful for producing a polypeptide in a Bacillus strain.
CC The present sequence is a PCR primer used in the construction of a
CC short consensus mutated Bacillus amyloliquefaciens alpha-amylase (amyQ)
CC dimer promoter gene expression cassette, used in the exemplification of
CC the invention.

XX Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 10

AAF62648
ID AAF62648 standard; DNA; 33 BP.

XX AC AAF62648;

DT 03-MAY-2001 (first entry)

XX amyQ promoter oligonucleotide #6.

DE Pectin acetyltransferase; degrade; plant cell wall; ss.

XX OS Synthetic.

XX PN US6184028-B1.

XX PD 06-FEB-2001.

XX PF 26-AUG-1999; 99US-0384305.

XX PR 26-AUG-1999; 99US-0384305.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PI Thomas MD, Brown KM;

XX DR WPI; 2001-190946/19.

XX Novel isolated polypeptide having pectin acetyltransferase activity useful
PT for degrading pectic substances and in degradation or modification of
PT acetylated pectins and plant cell walls -

XX Example 10; Column 29; 35pp; English.

CC The present invention relates to Bacillus subtilis pectin
CC acetyltransferase protein. The invention is useful for degrading
CC a pectic substance. It is also useful for degrading soluble and
CC insoluble pectins with varying degrees of esterification,
CC clarification etc. The protein may be used alone or in
CC combination with other enzymes for the degradation or modification
CC of acetylated pectins, degradation or modification of plant cell walls.

XX Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 11

AAV38595
ID AAV38595 standard; DNA; 34 BP.

XX AC AAV38595;

DT 17-SEP-1998 (first entry)

XX PCR primer 19 used to amplify a portion of the amyE gene.

XX amyE gene; strain Al64; mutant Bacillus; biosynthesis; surfactin;
 KW production; hormone; enzyme; receptor; PCR primer; ss.
 XX Synthetic.
 OS Bacillus subtilis.
 OS WO9822598-A1.
 PN 28-MAY-1998.
 XX 18-NOV-1997; 97WO-US21084.
 XX 12-JUN-1997; 97US-0049441.
 PR 18-NOV-1996; 96US-0749521.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PI Adams LF, Brown S, Sloma A, Sternberg D;
 XX WPI; 1998-312483/27.
 DR Producing polypeptides in Bacillus cells - which are modified to
 PT reduce production of surfactin, useful for, e.g. producing hormones
 PT or enzymes
 XX Example 5; Page 25; 43pp; English.
 PS PCR primers AAV38593-94 and AAV38595-96 were used to amplify the
 CC different portions of the amyE gene of Bacillus subtilis strain Al64.
 CC The portions were cloned to create a partially deleted amyE gene. This
 CC was used in the course of the invention. The specification describes the
 CC production of a polypeptide. This comprises culturing a mutant Bacillus
 CC that includes nucleic acid encoding the polypeptide and nucleic acid
 CC containing a modified form of at least 1 gene involved in biosynthesis
 CC of a surfactin or its isoforms. The mutant organism produces less
 CC surfactin, or isoform, than wild-type Bacillus. The method may be used
 CC to produce hormones, enzymes and receptors, which are either
 CC heterologous or homologous.
 XX Sequence 34 BP; 5 A; 10 C; 11 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTTAAGGCG 11
 Db 3 gccttaaggcg 13

RESULT 12
 AAV82401
 ID AAV82401 standard; DNA; 34 BP.
 XX AAV82401;
 AC 12-APR-1999 (first entry)
 XX Bacillus subtilis alpha-amylase amyE gene PCR primer 19.
 XX Alpha-amylase; amyE gene; protease; PCR; primer; ss.
 KW Synthetic.
 XX Bacillus subtilis.
 OS WO9856927-A2.
 PN 17-DEC-1998.
 XX 09-JUN-1998; 98WO-US12005.

PR 12-JUN-1997; 97US-0873479.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA Christianson L, Sloma A;
 PI WPI; 1999-080908/07.
 XX Novel protease from Bacillus subtilis LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing
 PT Example 5; Page 30; 77pp; English.
 PS Primers 19 and 20 (see AAV82401-02) are designed for use in the PCR
 CC amplification of a downstream portion (nucleotides 445-953,
 CC where the A of the ATG start codon is +1) of the alpha-amylase
 CC amyE gene from Bacillus subtilis Al64 chromosomal DNA. Primer 19
 CC includes SfiI and NotI sites. An upstream portion of the amyE gene
 CC (nucleotides -421 to +77) was amplified using primers 17 and 18
 CC (see AAV82399-400). The 2 fragments were then spliced by PCR using
 CC primers 17 and 20. A splicing by overlap extension (SOE) reaction
 CC was performed to create a deleted version of the amyE gene. Mutant
 CC host cells having a deleted spoIIAC gene, a deleted amyE gene and
 CC deleted protease genes have been constructed. These can be used in
 CC the production of heterologous polypeptides. The invention also
 CC relates to Bacillus proteases (see AAV89547-48) suitable for use in
 CC detergent compositions and in leather processing.
 XX Sequence 34 BP; 5 A; 10 C; 11 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTTAAGGCG 11
 Db 3 gccttaaggcg 13

RESULT 13
 AAV38594/C
 ID AAV38594 standard; DNA; 36 BP.
 XX AAV38594;
 AC 17-SEP-1998 (first entry)
 XX PCR primer 18 used to amplify a portion of the amyE gene.
 XX amyE gene; strain Al64; mutant Bacillus; biosynthesis; surfactin;
 KW production; hormone; enzyme; receptor; PCR primer; ss.
 OS Synthetic.
 OS Bacillus subtilis.
 XX WO9822598-A1.
 PN 28-MAY-1998.
 XX 18-NOV-1997; 97WO-US21084.
 XX 12-JUN-1997; 97US-0049441.
 PR 18-NOV-1996; 96US-0749521.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA Adams LF, Brown S, Sloma A, Sternberg D;
 PI WPI; 1998-312483/27.
 DR Producing polypeptides in Bacillus cells - which are modified to
 PT reduce production of surfactin, useful for, e.g. producing hormones

PT or enzymes
PS Example 5; Page 25; 43pp; English.
XX
XX PCR primers AAV38593-94 and AAV38595-96 were used to amplify the
CC different portions of the amyE gene of *Bacillus subtilis* strain Al64.
CC The portions were cloned to create a partially deleted amyE gene. This
CC was used in the course of the invention. The specification describes the
CC production of a polypeptide. This comprises culturing a mutant *Bacillus*
CC that includes nucleic acid encoding the polypeptide and nucleic acid
CC containing a modified form of at least 1 gene involved in biosynthesis
CC of a surfactin or its isoforms. The mutant organism produces less
CC surfactin, or isoform, than wild-type *Bacillus*. The method may be used
CC to produce hormones, enzymes and receptors, which are either
XX heterologous or homologous.
SQ Sequence 36 BP; 11 A; 12 C; 10 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 22 GCCTTAAGGCG 12

RESULT 14
AAV82400/C
ID AAV82400 standard; DNA; 36 BP.
XX
AC AAV82400;
XX
DT 12-APR-1999 (first entry)
XX
DE *Bacillus subtilis* alpha-amylase amyE gene PCR primer 18.
XX
KW Alpha-amylase; amyE gene; protease; PCR; primer; ss.
XX
OS Synthetic.
OS *Bacillus subtilis*.
XX
PN W09856927-A2.
XX
PD 17-DEC-1998.
XX
PF 09-JUN-1998; 98WO-US12005.
XX
PR 12-JUN-1997; 97US-0873479.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Christianson L, Sloma A;
XX
DR WPI; 1999-080908/07.
XX
PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
PT dishwashing detergents and for leather processing
PS Example 5; Page 29; 77pp; English.

Primers 17 and 18 (see AAV82399-400) are designed for use in the PCR
CC amplification of an upstream portion (nucleotides -421 to +77,
CC where the A of the ATG start codon is +1) of the alpha-amylase
CC amyE gene from *Bacillus subtilis* Al64 chromosomal DNA. Primer 18
CC (nucleotides 445-953) was amplified using primers 19 and 20 (see
CC AAV82401-02). The 2 fragments were then spliced by PCR using
CC primers 17 and 20. A splicing by overlap extension (SOE) reaction
CC was performed to create a deleted version of the amyE gene. Mutant
CC host cells having a deleted spoIIAC gene, a deleted amyE gene and
CC deleted protease genes have been constructed. These can be used in

CC the production of heterologous polypeptides. The invention also
CC relates to *Bacillus* proteases (see AAV89547-48) suitable for use in
CC detergent compositions and in leather processing.
XX
SQ Sequence 36 BP; 11 A; 12 C; 10 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 22 GCCTTAAGGCG 12

RESULT 15
AAZ23328
ID AAZ23328 standard; DNA; 37 BP.
XX
AC AAZ23328;
XX
DT 06-DEC-1999 (first entry)
XX
DE *Bacillus licheniformis* amyL gene promoter PCR primer 1.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone; amyL;
KW antibody; reporter; marker gene; cell regulation; serine protease;
KW SAVINASE; PCR primer; ss.
XX
OS Synthetic.
OS *Bacillus licheniformis*.
XX
PN W09943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
PR 26-FEB-1998; 98US-0031442.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 1999-561370/47.
XX
PT Production of polypeptide in *Bacillus* using specific promoters,
PT particularly for producing enzymes -
XX
PS Example 6; Page 26; 90pp; English.

This invention describes a novel method for the production of a
CC polypeptide in *Bacillus* using specific tandem or consensus promoters.
CC The method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a PCR primer used to amplify
CC the *Bacillus licheniformis* alpha-amylase amyL promoter which is used in
CC the construction of an amyL promoter/serine protease, SAVINASE (RTM) gene
CC expression cassette described in the method of the invention.
XX
SQ Sequence 37 BP; 4 A; 12 C; 10 G; 11 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 37;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
| | | | | | | | | |
Db 5 gccttaaggc 15

Search completed: January 17, 2002, 12:02:19
Job time: 17726 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:00 ; Search time 6501.33 seconds
(without alignments)
18.181 Million cell updates/sec

Title: US-09-242-202a-10_COPY_1_11
Perfect score: 11
Sequence: 1 GCCTTAAGGC 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST.*
1: em_estfun.*
2: em_esthum.*
3: em_estin.*
4: em_estom.*
5: em_estpl.*
6: em_estba.*
7: em_estro.*
8: em_estov.*
9: em_hic.*
10: gb_estl.*
11: gb_est2.*
12: gb_hic.*
13: gb_gss.*
14: em_gss_fun.*
15: em_gss_hum.*
16: em_gss_inv.*
17: em_gss_pln.*
18: em_gss_pro.*
19: em_gss_rod.*
20: em_gss_vrt.*
21: em_gss_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	100.0	100	AA093528	AA093528 c10250.se
2	11	100.0	115	A2119770	A2119770 RPI-23-4
3	11	100.0	130	AQ355689	AQ355689 CITBI-E1-
4	11	100.0	136	AA345723	AA345723 EST51785
5	11	100.0	140	AZ123745	AZ123745 RPI-23-4
6	11	100.0	147	N27010	N27010 yw73a09.s1
7	11	100.0	149	BE245846	BE245846 TCBAPIE19
8	11	100.0	158	N41203	N41203 HRI179-F.Adu
9	11	100.0	159	BE244007	BE244007 TCBAPIE04
10	11	100.0	165	AQ012949	AQ012949 CIT-HSP-2
11	11	100.0	172	AW308236	AW308236 3366 MARC
12	11	100.0	181	AI303597	AI303597 u181a12.x

c 13	11	100.0	182	10	AI800916
c 14	11	100.0	185	11	BF605530
c 15	11	100.0	199	13	AZ075637
c 16	11	100.0	203	10	AA303217
c 17	11	100.0	207	10	AA498912
c 18	11	100.0	208	10	BE698384
c 19	11	100.0	213	11	F18421
c 20	11	100.0	214	10	AW197056
c 21	11	100.0	223	10	BB143920
c 22	11	100.0	227	10	AV321312
c 23	11	100.0	230	11	F25526
c 24	11	100.0	231	10	AW920467
c 25	11	100.0	231	11	N23615
c 26	11	100.0	232	11	BG552774
c 27	11	100.0	232	11	BI262529
c 28	11	100.0	234	10	BB152224
c 29	11	100.0	235	10	AI306150
c 30	11	100.0	241	10	AV062684
c 31	11	100.0	241	10	AV146659
c 32	11	100.0	243	11	Z38396
c 33	11	100.0	246	10	AA094468
c 34	11	100.0	251	10	AU127469
c 35	11	100.0	251	13	AZ366083
c 36	11	100.0	252	10	AV649344
c 37	11	100.0	252	10	BB606444
c 38	11	100.0	254	10	AA576933
c 39	11	100.0	255	11	N83262
c 40	11	100.0	255	11	BF60539
c 41	11	100.0	256	10	AV281033
c 42	11	100.0	257	11	BF509446
c 43	11	100.0	259	10	AI350354
c 44	11	100.0	259	10	BB390531
c 45	11	100.0	259	13	AZ217098

ALIGNMENTS

RESULT 1

AA093528	AA093528	100 bp	mRNA	EST	25-OCT-1996
LOCUS	c10250.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens				
DEFINITION	cDNA 5', mRNA sequence.				
ACCESSION	AA093528				
VERSION	AA093528.1	GI:1639161			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Liew,C.C.				
TITLE	CDNAs from fetal heart (1996)				
JOURNAL	Unpublished (1996)				
COMMENT	Contact: Liew CC Department of Laboratory Medicine and Pathobiology University of Toronto Banting Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel: 4169788758 Fax: 4169785650 Email: liewccc@utcc.utoronto.ca				
PCR Primers					
FORWARD:	5' GCCAAGCTCGAATTAACCTCCTACTAAGGG 3'				
BACKWARD:	5' CCAGTGAATGTAACTACGACTACTATAGGCG 3'				
Seq primer:	5' GAAATTAACCTCCTACTAAGGG 3'				

FEATURES

Location/Qualifiers
1..100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2: XhoI; mRNA was purified from human fetal hearts (8-10

weeks). cDNA was synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

BASE COUNT

10 a 39 c 27 g 24 t

Query Match

Best Local Similarity 100.0%; Score 11; DB 10; Length 100;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11

Db 78 GCCTTAAGGCG 88

RESULT 2

AZ119770/c
LOCUS AZ119770 115 bp DNA GSS 12-MAY-2000
DEFINITION RPCI-23-479A23-TV RPCI-23 Mus musculus genomic clone RPCI-23-479A23
, DNA sequence.

ACCESSION AZ119770

VERSION AZ119770.1 GI:7785028

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 115)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroi, M., de Jong, P.
and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Other_GSSs: RPCI-23-479A23.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 479 row: A column: 23

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..115

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-479A23"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

35 a 32 c 28 g 20 t

ORIGIN

Query Match

100.0%; Score 11; DB 13; Length 115;

Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11

Db 114 GCCTTAAGGCG 104

RESULT 3

AQ355689

LOCUS

DEFINITION AQ355689 130 bp DNA GSS 24-JAN-1999
CITBI-E1-2538F8.TR CITBI-E1 Homo sapiens genomic clone 2538F8, DNA

sequence.

ACCESSION AQ355689

VERSION AQ355689.1 GI:4182862

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 130)

AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and

Venter, J.C.

TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready

Map Building

Unpublished (1997)

CONTACT: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..130

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="2538F8"

/clone_lib="CITBI-E1"

/sex="male"

/cell_type="sperm"

/note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;

Caltech Human BAC Library D"

42 a 23 c 21 g 44 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 11; DB 13; Length 130;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11

Db 23 GCCTTAAGGCG 33

RESULT 4

AA345723

LOCUS

DEFINITION AA345723 136 bp mRNA EST 21-APR-1997
EST51785 Gall bladder I Homo sapiens cDNA 5' end similar to similar

to prollylcarboxypeptidase, mRNA sequence.

ACCESSION AA345723

VERSION AA345723.1 GI:1997971

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 136)
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Pannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.
 TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi.html>)
 Seq primer: M13 Reverse

FEATURES
 source
 1. .136
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):147260"
 /db_xref="taxon:9606"
 /clone_lib="Gall bladder I"
 /sex="female"
 /dev_stage="adult, 25 yrs"
 /note="Organ: gall bladder; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 18 a 59 c 30 g 25 t 4 others
 ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 136;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 |||||
 Db 89 GCCTTAAGGCG 99

RESULT 5
 AZ123745/c 140 bp DNA GSS 12-MAY-2000
 LOCUS
 DEFINITION RPCI-23-479M23-TV RPCI-23 Mus musculus genomic clone RPCI-23-479M23
 DNA sequence.
 ACCESSION AZ123745
 VERSION AZ123745.1 GI:7792734
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 140)
 REFERENCE
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
 B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
 and Fraser,C.M.
 TITLE Mouse BAC End Sequences from Library RPCI-23

JOURNAL
 COMMENT Unpublished (1999)
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 plate: 479 row: M column: 23
 Seq primer: T7
 Class: BAC ends.

FEATURES
 Location/Qualifiers
 source
 1. .140
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-479M23"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:
 EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBACE3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."
 BASE COUNT 45 a 39 c 34 g 22 t
 ORIGIN

Query Match 100.0%; Score 11; DB 13; Length 140;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 |||||
 Db 139 GCCTTAAGGCG 129

RESULT 6
 N27010 147 bp mRNA EST 29-DEC-1995
 LOCUS yw73a09.s1 Soares.placenta.8to9weeks_2NbHP8to9W Homo sapiens cDNA
 DEFINITION Clone IMAGE:257848 3' similar to gb:M30142 COMPLEMENT
 DECAT-ACCELERATING FACTOR 1 PRECURSOR (HUMAN);, mRNA sequence.
 N27010
 VERSION N27010.1 GI:1141358
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 147)
 REFERENCE
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
 M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
 R., Williamson,A., Wohlmann,P. and Wilson,R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Willson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence starts: 1

High quality sequence stops: 1
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace Consortium (info@image.llnl.gov) for further information.
 Possible reversed clone: polyt not found
 Seq primer: m13 -40 forward
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 147

FEATURES

source

/organism="Homo sapiens"
 /db_xref="GDB:3887458"
 /db_xref="taxon:9606"
 /clone="IMAGE:257848"
 /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M. Fatima Bonaldo."
 41 a 26 c 33 g 36 t 11 others

BASE COUNT

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 147;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTTAAGGCG 11

Db 102 GCCTTAAGGCG 112

RESULT 7

BE245846/c
 LOCUS
 DEFINITION TCBAP1E1938 Pediatric pre-B cell acute lymphoblastic leukemia
 Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP1938, mRNA
 sequence.
 BE245846
 BE245846.1 GI:9097594
 EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 149)
 REFERENCE
 AUTHORS Wei, Y., Tsang, Y. T. M., Mei, G., Ku, J. M., Ali-Osman Jr., F. R., Muzny, D.,
 Bouck, J., Gibbs, R. A. and Margolin, J. F.
 TITLE Pediatric Leukemia cDNA Sequencing Project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@xccc.org
 Seq primer: M13 primer
 Location/Qualifiers
 1. 149

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCBAP1938"
 /clone_lib="Pediatric pre-B cell acute lymphoblastic
 leukemia Baylor-HGSC project-TCBA"

/sex="male"
 /tissue_type="leukopheresis"
 /cell_type="pre-B cell"
 /dev_stage="pediatric 2 years"
 /lab_host="DH10B"

/note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
 First strand cDNA was primed with an anchored
 XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCAGGAG(T)VN
 3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second
 strand was primed with a BamHI-dc primer
 [5'AGAGAGCTCGATCCGCGCGCAATAATAAT(C) 3'].
 Double-stranded cDNA was then digested with BamHI and
 XhoI and directionally cloned into the BamHI and SalI
 sites of lambda pSB vector. Library went through one
 round of normalization. Library was constructed by Wei
 Yu."

BASE COUNT 35 a 56 c 46 g 10 t 2 others

ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 149;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTTAAGGCG 11

Db 50 GCCTTAAGGCG 40

RESULT 8

N41203/c
 LOCUS
 DEFINITION HRI79-F Adult heart, Clontech Homo sapiens cDNA clone RI79-F, mRNA
 sequence.
 N41203
 N41203.1 GI:1164801
 EST.
 KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 158)
 REFERENCE
 AUTHORS Waye, M. Y., Cheung, H. K. Y., Lam, W. Y., Law, P. T. W., Lo, A. S. Y., Lui,
 V. W. Y., Luk, S. C. W., Tsui, S. K. W., Tung, C. K. C., Yan, N. Y. H., Liew,
 C. C. and Lee, C. Y.
 TITLE Gene expression of adult human heart as revealed by random
 sequencing of cDNA library
 JOURNAL Miami Winter Biotechnol. Symp. Proc. 6, 90 (1995)
 COMMENT Contact: Wayne Mary M.Y.
 Department of Biochemistry
 The Chinese University of Hong Kong
 Rm 302C, Basic Medical Science Building, The Chinese University of
 Hong Kong, Shatin, N.T., Hong Kong.
 Tel: 8526096874
 Fax: 8526035123
 Email: bl33723@vax.csc.cuhk.hk
 Seq primer: GTGGCGGAGCTCTCTGGAGCC.
 Location/Qualifiers
 1. 158

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RI79-F"
 /clone_lib="Adult heart, Clontech"
 /lab_host="E. coli Y1090"
 /note="Vector: Lambda gtl1; Site_1: EcoRI; Site_2: EcoRI"

BASE COUNT

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 158;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 a 50 c 35 g 24 t

ORIGIN

```

Qy 1 GCCTTAAGGC 11
Db 141 GCCTTAAGGC 131

RESULT 9
BE244007/c
LOCUS
DEFINITION
TCBAP1E0434 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP0434, mRNA
sequence.
ACCESSION
BE244007
VERSION
BE244007.1 GI:9095747
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 159)
AUTHORS
Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R., Muzny, D.,
Bouck, J., Gibbs, R.A., and Margolin, J.F.
PEDIATRIC Leukemia cDNA Sequencing Project
Unpublished (2000)
TITLE
Contact: Dr. Judith F. Margolin
JOURNAL
Texas Children's Cancer Center and Human Genome Sequencing Center
COMMENT
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@ccc.org
Seq primer: M13 primer.
FEATURES
source
1..159
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TCBAP0434"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project-TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH108"
/notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second
strand was primed with a BamHI-dC primer
[5'AGAGCTCGGATCGCGCGCGCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and
XhoI and directionally cloned into the BamHI and SalI
sites of lambda pSB vector. Library went through one
round of normalization. Library was constructed by Wei
Yu"
BASE COUNT 48 a 48 c 47 g 12 t 4 others
ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTTAAGGC 11
Db 60 GCCTTAAGGC 70

RESULT 11
AW308236/c
LOCUS
DEFINITION
3366 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION
AW308236
VERSION
AW308236.1 GI:6720599
KEYWORDS
EST.
SOURCE
cow.
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 172)
AUTHORS
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL
Genome Res. 11 (4), 626-630 (2001)
MEDLINE
21180013
COMMENT
Contact: Smith TPL

Qy 1 GCCTTAAGGC 11
Db 42 GCCTTAAGGC 32

RESULT 10
AQ012949
LOCUS
DEFINITION
CIT-HSP-2300F24.TF CIT-HSP Homo sapiens genomic clone 2300F24, DNA
sequence.
ACCESSION
AQ012949

```

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred
v0.960904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACACGATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 147 row: J column: 11
Seq primer: ATTAGTGCACATATAG.

FEATURES

source

Location/Qualifiers
1. .172
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC lBOV"
/tissue_type="pooled"
/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

40 a 61 c 47 g 24 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 172;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
|||||
DB 22 GCCTTAAGGCG 12

RESULT 12

AI303597

LOCUS

AI303597 181 bp mRNA EST 08-DEC-1998
ui81a12.x1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1888798 3' similar to O35533 035533 CARBOXYLESTERASE
PRECURSOR ;, mRNA sequence.

ACCESSION AI303597
VERSION AI303597
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 181)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL

COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:973122

Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
1. .181

/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888798"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
(CACTGTGTG); Site_2: DraIII (CACATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGCGCCTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5Kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTATAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGACA."
38 a 50 c 47 g 46 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 181;
Best Local Similarity 100.0%; Pred. No. 7.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
|||||
DB 81 GCCTTAAGGCG 91

RESULT 13

AI800916/c

LOCUS

AI800916 182 bp mRNA EST 19-DEC-1999
wg14d07.x1 Soares NSF F8_9W_OT_PA_P-S1 Homo sapiens cDNA clone
IMAGE:2365069 3' similar to gb:U10320 FRUCTOSE-1,6-BISPHOSPHATASE
(HUMAN);, mRNA sequence.

ACCESSION AI800916
VERSION AI800916
KEYWORDS EST.
SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 182)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Insert Length: 1617

Seq primer: -400P from Gibco

High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
1. .182
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2365069"
/clone_lib="Soares NSF_F8_9W_OT_PA_P-S1"
/lab_host="DH10B"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer
in a subtractive hybridization reaction. The driver was

PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-149103, 148872-149255, 15002 - 150407, 151176-153327 Soares NB2HF8-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 35 a 56 c 51 g 40 t

ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
|||||

Db 85 GCCTTAAGGC 75

RESULT 14

LOCUS BF605530/c

DEFINITION 271955 MARC 3BOV Bos taurus cDNA 5', mRNA EST 25-APR-2001

ACCESSION BF605530

VERSION BF605530.1 GI:11705237

KEYWORDS EST.

SOURCE COW.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W., Rohrer, G.A., Chitko-McKown, C.G., Pertea, G., Holt, I., Karamecheva, S., Liang, F., Quackenbush, J. and Keele, J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 55 row: F column: 24

Seq primer: ATTTAGTGACACTATAG.

FEATURES

source

1. .185

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 51 a 39 c 48 g 47 t

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 185;

Best Local Similarity 100.0%; Pred. No. 7.4e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11

|||||

Db 137 GCCTTAAGGC 127

RESULT 15

LOCUS AZ075637/c

DEFINITION RPCI-23-408L19.TV RPCI-23 Mus musculus genomic clone RPCI-23-408L19 DNA sequence.

ACCESSION AZ075637

VERSION AZ075637.1 GI:7368534

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 199)

AUTHORS Zhao, S., Niernman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong

(pieter@jeong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (inforesgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 408 row: L column: 19

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. .199

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-408L19"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 53 a 52 c 33 g 61 t

ORIGIN

Query Match 100.0%; Score 11; DB 13; Length 199;

Best Local Similarity 100.0%; Pred. No. 7.4e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11

|||||

Db 36 GCCTTAAGGC 26

Search completed: January 17, 2002, 08:54:03
Job time: 7115 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:48:13 ; Search time 10436.3 Seconds
(without alignments)
2252.560 Million cell updates/sec

Title: US-09-242-202A-16

Perfect score: 1425

Sequence: 1 TCCCATGGCGGCGATTCTTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: gb_ba.*	GenEmbl.*
2: gb_htg.*	
3: gb_in.*	
4: gb_om.*	
5: gb_ov.*	
6: gb_pat.*	
7: gb_ph.*	
8: gb_pl.*	
9: gb_pr.*	
10: gb_ro.*	
11: gb_sts.*	
12: gb_sy.*	
13: gb_un.*	
14: gb_vi.*	
15: em_ba.*	
16: em_fun.*	
17: em_hum.*	
18: em_in.*	
19: em_om.*	
20: em_or.*	
21: em_ov.*	
22: em_pat.*	
23: em_ph.*	
24: em_pl.*	
25: em_ro.*	
26: em_sts.*	
27: em_sy.*	
28: em_un.*	
29: em_vi.*	
30: em_htgo_hum.*	
31: em_htgo_inv.*	
32: em_htgo_rod.*	
33: em_htgo_hum.*	
34: em_htg_inv.*	
35: em_htg_rod.*	
36: em_htg_other.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	454.8	31.9	11795	6	AX027785	AX027785 Sequence
c 2	454.8	31.9	13254	6	AR038307	AR038307 Sequence
c 3	454.8	31.9	13254	6	AR038321	AR038321 Sequence
c 4	454.8	31.9	13254	6	I58596	I58596 Sequence 15
c 5	454.8	31.9	13254	6	I58610	I58610 Sequence 17
c 6	454.2	31.9	6649	6	AX180726	AX180726 Sequence
c 7	453.8	31.8	1905	6	AR027070	AR027070 Sequence
c 8	453.8	31.8	1905	6	I86203	I86203 Sequence 9
c 9	453.8	31.8	3273	12	ARP327	V000803 Artificial
c 10	453.8	31.8	4023	6	A49702	A49702 Sequence 8
c 11	453.8	31.8	4060	12	XXU13846	U13846 pCM7 clonin
c 12	453.8	31.8	4150	12	SYNBR329	J01753 Plasmid pBR
c 13	453.8	31.8	4410	6	A52326	A52326 Sequence 1
c 14	453.8	31.8	4410	6	AR027062	AR027062 Sequence
c 15	453.8	31.8	4410	6	I86195	I86195 Sequence 1
c 16	453.8	31.8	5241	6	A49695	A49695 Sequence 1
c 17	453.8	31.8	5241	6	A49696	A49696 Sequence 2
c 18	453.8	31.8	5241	6	A77043	A77043 Sequence 1
c 19	453.8	31.8	5241	6	A77044	A77044 Sequence 2
c 20	453.8	31.8	5365	6	I09267	I09267 Sequence 34
c 21	453.8	31.8	5413	6	I09270	I09270 Sequence 37
c 22	453.8	31.8	5518	6	I09268	I09268 Sequence 35
c 23	453.8	31.8	5566	6	I09269	I09269 Sequence 36
c 24	453.8	31.8	5629	6	A08457	A08457 pHS 164 mRN
c 25	453.8	31.8	6149	6	I09252	I09252 Sequence 19
c 26	453.8	31.8	6151	6	I09251	I09251 Sequence 18
c 27	453.8	31.8	6477	6	A08458	A08458 pinf 4-49 m
c 28	453.8	31.8	7218	12	ASPNGV1	X99274 Artificial
c 29	453.8	31.8	7252	6	AX003206	AX003206 Sequence
c 30	453.8	31.8	8119	6	AR002333	AR002333 Sequence
c 31	453.8	31.8	8119	6	AR018032	AR018032 Sequence
c 32	453.8	31.8	8119	6	AR055698	AR055698 Sequence
c 33	453	31.8	1968	12	AF087567	AF087567 Cloning v
c 34	453	31.8	1969	12	AF086840	AF086840 Cloning v
c 35	453	31.8	1969	12	AF086841	AF086841 Cloning v
c 36	453	31.8	1969	12	AF086842	AF086842 Cloning v
c 37	453	31.8	1969	12	AF086843	AF086843 Cloning v
c 38	453	31.8	1969	12	AF086844	AF086844 Cloning v
c 39	453	31.8	1969	12	AF086845	AF086845 Cloning v
c 40	453	31.8	1969	12	AF086846	AF086846 Cloning v
c 41	453	31.8	1969	12	AF086847	AF086847 Cloning v
c 42	453	31.8	1969	12	AF086848	AF086848 Cloning v
c 43	453	31.8	1969	12	AF086849	AF086849 Cloning v
c 44	453	31.8	1969	12	AF086850	AF086850 Cloning v
c 45	453	31.8	1969	12	AF086851	AF086851 Cloning v

ALIGNMENTS

RESULT 1	AX027785/c	AX027785	11795 bp	DNA	PAT	16-SEP-2000
LOCUS	Sequence 4 from Patent WO0043527.					
DEFINITION	AX027785					
ACCESSION	AX027785.1					
VERSION	GI:10188633					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

synthetic construct.
synthetic construct.
artificial sequence.
1 (bases 1 to 11795)
Bollen,A., Jacobs,P., Jacquet,A., Haumont,M. and Massaer,M.G.
Varicella-zoster virus vaccines
Patent: WO 0043527-A 4 27-JUL-2000;
SMITHKLINE BEECHAM BIOLOG (BE) ; BOLLEN ALEX (BE) ; JACOBS PAUL (BE) ; JACQUET ALAIN (BE) ; HAUMONT MICHELE (BE) ; MASSAER MARC (BE) ;
GEORGES FRANCIS (BE)
Location/Qualifiers
1. .11795
/organism="synthetic construct"

```
/db_xref="taxon:32630"
BASE COUNT 2917 a 2760 c 2957 g 3161 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 11795;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 986 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 927
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGATAGTCTGTCTTACCGGGTTG 831
DB 926 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGATAGTCTGTCTTACCGGGTTG 867
QY 832 GACTCAAGACCATATTACCGGATAAGGCGGAGCGGTGCGGCTGAAGCGGGGTTGTCGTG 891
DB 866 GACTCAAGACCATATTACCGGATAAGGCGGAGCGGTGCGGCTGAAGCGGGGTTGTCGTG 807
QY 892 ACACAGCCCGCTTGGAGCGAAGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 806 ACACAGCCCGCTTGGAGCGAAGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 747
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGAGGATATCCGGTAAAGCGGCAGG 1011
DB 746 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGAGGATATCCGGTAAAGCGGCAGG 687
QY 1012 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAACGCGTGTATCTTTATAGT 1071
DB 586 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAACGCGTGTATCTTTATAGT 627
QY 1072 CCTGTGGGTTTCCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
DB 626 CCTGTGGGTTTCCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 567
QY 1132 CGGAGCCTATGGAAGAAAGCGGACGAGCGGCGGGGG 1169
DB 566 CGGAGCCTATGGAAGAAAGCGGACGAGCGGCGGGGG 529
RESULT 2
AR038307/c 29-SEP-1999
LOCUS AR038307 13254 bp DNA PAT
DEFINITION Sequence 156 from patent US 5804440.
ACCESSION AR038307
VERSION AR038307.1 GI:5957024
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5804440-A 156 08-SEP-1998;
FEATURES
Location/Qualifiers
source 1..13254
BASE COUNT 3206 a 3559 c 3251 g 3238 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 6449 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 6390
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
/db_xref="taxon:32630"
BASE COUNT 2917 a 2760 c 2957 g 3161 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 11795;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 986 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 927
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGATAGTCTGTCTTACCGGGTTG 831
DB 926 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGATAGTCTGTCTTACCGGGTTG 867
QY 832 GACTCAAGACCATATTACCGGATAAGGCGGAGCGGTGCGGCTGAAGCGGGGTTGTCGTG 891
DB 866 GACTCAAGACCATATTACCGGATAAGGCGGAGCGGTGCGGCTGAAGCGGGGTTGTCGTG 807
QY 892 ACACAGCCCGCTTGGAGCGAAGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 806 ACACAGCCCGCTTGGAGCGAAGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 747
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGAGGATATCCGGTAAAGCGGCAGG 1011
DB 746 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGAGGATATCCGGTAAAGCGGCAGG 687
QY 1012 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAACGCGTGTATCTTTATAGT 1071
DB 586 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAACGCGTGTATCTTTATAGT 627
QY 1072 CCTGTGGGTTTCCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
DB 626 CCTGTGGGTTTCCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 567
QY 1132 CGGAGCCTATGGAAGAAAGCGGACGAGCGGCGGGGG 1169
DB 566 CGGAGCCTATGGAAGAAAGCGGACGAGCGGCGGGGG 529
RESULT 3
AR038321 29-SEP-1999
LOCUS AR038321 13254 bp DNA PAT
DEFINITION Sequence 170 from patent US 5804440.
ACCESSION AR038321
VERSION AR038321.1 GI:5957038
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5804440-A 170 08-SEP-1998;
FEATURES
Location/Qualifiers
source 1..13254
BASE COUNT 3238 a 3251 c 3559 g 3206 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 6806 GTGTAGCGTGTAGGCGCCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 6865
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
DB 6866 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 6925
QY 832 GACTCAAGACCATATTACCGGATAAGGCGGAGCGGCTGAAGCGGGGTTGTCGTG 891
DB 6926 GACTCAAGACCATATTACCGGATAAGGCGGAGCGGCTGAAGCGGGGTTGTCGTG 6985
QY 892 ACACAGCCCGCTTGGAGCGAAGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 6986 ACACAGCCCGCTTGGAGCGAAGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 7045
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGAGGATATCCGGTAAAGCGGCAGG 1011
DB 7046 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGAGGATATCCGGTAAAGCGGCAGG 7105
```

QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 1071
|||||
Db 7106 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 7165
QY 1072 CCGTGTGGGTTTCCGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
|||||
Db 7166 CCGTGTGGGTTTCCGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 7225
QY 1132 CGGAGCCTATGGAAACGCGACGACGCGCGCGGGG 1169
|||||
Db 7226 CGGAGCCTATGGAAACGCGACGACGCGCGCGGGG 7263
RESULT 4
LOCUS I58596/c I58596 13254 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 156 from patent US 5652138.
ACCESSION I58596
VERSION I58596.1 GI:2477834
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5652138-A 156 29-JUL-1997;
FEATURES Location/Qualifiers
source 1..13254
BASE COUNT 3206 a 3559 c 3251 g 3238 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACCTCTAGCAGCGCTACATACCTCGCT 771
|||||
Db 6449 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACCTCTAGCAGCGCTACATACCTCGCT 6390
QY 772 CTGCTAATCTGTACCAGTGGCTGCTCCAGTGGCGATAGTCTCTTACCGGGTTG 831
|||||
Db 6389 CTGCTAATCTGTACCAGTGGCTGCTCCAGTGGCGATAGTCTCTTACCGGGTTG 6330
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAACGGGGGTTCTGTC 891
|||||
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAACGGGGGTTCTGTC 6270
QY 892 ACACAGCCCGAGTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 951
|||||
Db 6269 ACACAGCCCGAGTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCCGACGCTTCCCGAGGAGAAAGCGGAGTATCCGTTAAGCGGCAGG 1011
|||||
Db 6209 TGAGAAAGCCGACGCTTCCCGAGGAGAAAGCGGAGTATCCGTTAAGCGGCAGG 6150
QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 1071
|||||
Db 6149 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 6090
QY 1072 CCGTGTGGGTTTCCGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
|||||
Db 6089 CCGTGTGGGTTTCCGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 6030
QY 1132 CGGAGCCTATGGAAACGCGACGACGCGCGCGGGG 1169
|||||
Db 6029 CGGAGCCTATGGAAACGCGACGACGCGCGCGGGG 5992

RESULT 5
LOCUS I58610 I58610 13254 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 170 from patent US 5652138.
ACCESSION I58610
VERSION I58610.1 GI:2477848
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5652138-A 170 29-JUL-1997;
FEATURES Location/Qualifiers
source 1..13254
BASE COUNT 3238 a 3251 c 3559 g 3206 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACCTCTAGCAGCGCTACATACCTCGCT 771
|||||
Db 6806 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACCTCTAGCAGCGCTACATACCTCGCT 6865
QY 772 CTGCTAATCTGTACCAGTGGCTGCTCCAGTGGCGATAGTCTCTTACCGGGTTG 831
|||||
Db 6866 CTGCTAATCTGTACCAGTGGCTGCTCCAGTGGCGATAGTCTCTTACCGGGTTG 6925
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAACGGGGGTTCTGTC 891
|||||
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAACGGGGGTTCTGTC 6985
QY 892 ACACAGCCCGAGTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 951
|||||
Db 6986 ACACAGCCCGAGTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 7045
QY 952 TGAGAAAGCCGACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
|||||
Db 7046 TGAGAAAGCCGACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 7105
QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 1071
|||||
Db 7106 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 7165
QY 1072 CCGTGTGGGTTTCCGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
|||||
Db 7166 CCGTGTGGGTTTCCGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 7225
QY 1132 CGGAGCCTATGGAAACGCGACGACGCGCGCGGGG 1169
|||||
Db 7226 CGGAGCCTATGGAAACGCGACGACGCGCGCGGGG 7263
RESULT 6
LOCUS AX180726/c AX180726 6649 bp DNA PAT 06-AUG-2001
DEFINITION Sequence 6 from Patent WO0146694.
ACCESSION AX180726
VERSION AX180726.1 GI:15132581
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 6649)
AUTHORS Joly,E.
TITLE A bioluminescence resonance energy transfer (bret) fusion molecule and method of use

Db	2034	CTGCTAATCCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG	1975
QY	832	GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGTGAACGGGGGTTCTGTCG	891
Db	1974	GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGTGAACGGGGGTTCTGTCG	1915
QY	892	ACACAGCCACGCTTGAGCGAAGCAGCTACACGGAACCTAGATACCTACAGCGTGAGCAT	951
Db	1914	ACACAGCCACGCTTGAGCGAAGCAGCTACACGGAACCTAGATACCTACAGCGTGAGCAT	1855
QY	952	TGAGAAAGCCACGCTTCCCGAAGGAGAAAGGCGCAGAGTATCCGCTAAGCGCGCAGG	1011
Db	1854	TGAGAAAGCCACGCTTCCCGAAGGAGAAAGGCGCAGAGTATCCGCTAAGCGCGCAGG	1795
QY	1012	GTCGGAACGAGGACGCGACGAGGAGCTTCCAGGGGGAACCGCTTATCTTTATAGT	1071
Db	1794	GTCGGAACGAGGACGCGACGAGGAGCTTCCAGGGGGAACCGCTTATCTTTATAGT	1735
QY	1072	CTGTGCGGTTTGGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTGTCGAGGGGG	1131
Db	1734	CTGTGCGGTTTGGCACCTCTGACTTGAGCGTCGATTTTGTGATGCTGTCGAGGGGG	1675
QY	1132	CGGAGCCTATGGAAGAACCCAGCAACGCGCGCGGGG	1168
Db	1674	CGGAGCCTATGGAAGAACCCAGCAACGCGCGCGGGG	1638
RESULT	11		
XXU13846/c			
LOCUS	XXU13846	4060 bp	DNA circular SYN 13-DEC-1994
DEFINITION	pcM7 cloning vector, complete sequence.		
ACCESSION	U13846		
VERSION	U13846.1	GI:595696	
KEYWORDS	chloramphenicol acetyltransferase; promoter analysis; beta-lactamase.		
SOURCE	unidentified cloning vector.		
ORGANISM	unidentified cloning vector.		
REFERENCE	artificial sequence; vectors.		
AUTHORS	1 (bases 1 to 4060)		
TITLE	Malone, J.A.		
JOURNAL	pcM7: A promoter analysis, chloramphenicol acetyltransferase gene		
REFERENCE	cartridge vector		
AUTHORS	Unpublished (1994)		
TITLE	2 (bases 1 to 4060)		
JOURNAL	Close, T.J. and Rodriguez, R.L.		
REFERENCE	Construction and characterization of the chloramphenicol-resistance		
AUTHORS	gene cartridge: a new approach to the transcriptional mapping of		
TITLE	extrachromosomal elements		
JOURNAL	Gene 20 (2), 305-316 (1982)		
MEDLINE	83158767		
REFERENCE	3 (bases 38 to 812)		
AUTHORS	Alton, N.K. and Vapnek, D.		
TITLE	Nucleotide sequence analysis of the chloramphenicol resistance		
JOURNAL	transposon Tn9		
MEDLINE	Nature 282 (5741), 864-869 (1979)		
REFERENCE	80078150		
AUTHORS	4 (bases 1 to 4060)		
TITLE	Malone, J.A.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (19-AUG-1994) James A. Malone, International Technical		
AUTHORS	Services, Molecular Biology Reagents Division, Pharmacia Biotech		
TITLE	Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA		
JOURNAL	Location/Qualifiers		
FEATURES			
source	1..4060		
gene	/organism="unidentified cloning vector"		
CDS	/db_xref="taxon:45196"		
	/lab_host="Escherichia coli"		
	67..726		
	/gene="cat"		
	67..726		
	/gene="cat"		
	/note="The cat gene in pcM7 is promoterless. The cat gene		

cartridge can be exised as a HindIII fragment"									
/citation=[2]									
/citation=[3]									
/codon_start=1									
/transl_table=11									
/product="chloramphenicol acetyltransferase"									
/protein_id="AAA57080.1"									
/db_xref="GI:595697"									
/translation="MEKKITGYTTVDISQSHRKEHFEAFQSAQCTYNQTVQLDITAF LTKVKNKHKYPAP IHTLARLMAHPEFRMAKDGELIVTWSVHPCYTVFHEQTFF SSLWEYHDDFROFLHIYSODVACYGENLAVFKPGFTENMFFVSANPVWSETFEDLNV ANMDDNEFAPVFTMGKYYTQGDVKLPLAIVQHVAVDGDFHVGRLMNLQYCYDEWQGG A"									
rep_origin	Complement(2233) /note="plasmid origin of replication; base 2233 represents the first base of the newly synthesized strand"								
gene	/direction=left complement(2992..3922)								
CDS	/gene="bla" complement(2992..3852) /gene="bla" /codon_start=1 /transl_table=11 /product="beta-lactamase" /protein_id="AAA57081.1" /db_xref="GI:595698" /translation="MSIOHFRVALIPFFAFLPVAFAHPETLVKVKDAEDQLGARVGY IELDNLSKILSEPRPEPMFTSKVLKCGAVLSVDQAGQQLGRRIRHYSQNDLVE YSPYETKLTDMGTIVRELCSAAITMSDNTAAILLATIGPKELTAFLLHMGDHVRL DRWPELNEATIPNDERGTTPAAAMATLRKLTLGELLTLASROQLDWMEDAKVAGPL LRSLAPAGWFTADKSGAGERSGRIIAALGPDGKPSRIVIYTTGSAQTMDERNRQIA EIGASLLIKHW" complement(3894..3922) /gene="bla"								
promoter	complement(3894..3922) /gene="bla"								
BASE COUNT	935	a	1078	c	1018	g	1029	t	
ORIGIN									
Query Match	31.8%; Score 453.8; DB 12; Length 4060;								
Best Local Similarity	99.6%; Pred. No. 6.5e-117;								
Matches	455;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	712	GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT	771						
Db	2667	GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT	2608						
QY	772	CTGCTAACTCTGTACCAAGTGGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG	831						
Db	2607	CTGCTAACTCTGTACCAAGTGGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG	2548						
QY	832	GACTCAAGACCATAGTTACCGGATAAGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTC	891						
Db	2547	GACTCAAGACCATAGTTACCGGATAAGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTC	2488						
QY	892	ACACAGCCCGCTTGGAGCGAAGCACTACACCGAACTAGATACCTACACGCTGAGCAT	951						
Db	2487	ACACAGCCCGCTTGGAGCGAAGCACTACACCGAACTAGATACCTACACGCTGAGCAT	2428						
QY	952	TGAGAAAGCGCCAGCTTCCGGAAGGGAGAAAGCGCACAGGTATCCGGTAAGCGGCAGG	1011						
Db	2427	TGAGAAAGCGCCAGCTTCCGGAAGGGAGAAAGCGCACAGGTATCCGGTAAGCGGCAGG	2368						
QY	1012	GTCCGAACAGGAGAGCGACGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGT	1071						
Db	2367	GTCCGAACAGGAGAGCGACGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGT	2308						
QY	1072	CCTGTGGGTTTCCCAACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCTGCAGGGGG	1131						
Db	2307	CCTGTGGGTTTCCCAACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCTGCAGGGGG	2248						
QY	1132	CGGAGCCTATGGAANAACGCCAAGCGCGCGCGGG	1168						
Db	2247	CGGAGCCTATGGAANAACGCCAAGCGCGCGCGCGAG	2211						

RESULT 12
SYNPBR329/c
LOCUS SYNPBR329 4150 bp DNA circular SYN 27-APR-1993
DEFINITION Plasmid pBR329, complete genome.
ACCESSION J01753
VERSION J01753.1 GI:208964
KEYWORDS cloning vector; complete genome.
SOURCE plasmid pbr329 from E.coli k-12.
ORGANISM unidentified cloning vector
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4150)
AUTHORS Covarrubias,L. and Bolivar,F.
TITLE Construction and characterization of new cloning vehicles: Vi.
Plasmid pBR329, a new derivative of pBR328 lacking the
482-base-pair inverted duplication
JOURNAL Gene 17, 79-89 (1982)
MEDLINE 8218782
COMMENT [1] only sequenced bases 160-340 and 3440-3560. the other regions
are from previously published sequences for pBR327 and Tn981 (see
fig. 5 for refs).
FEATURES
source Location/Qualifiers
1. .4150
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
BASE COUNT 970 a 1121 c 1021 g 1038 t
ORIGIN
Query Match 31.8%; Score 453.8; DB 12; Length 4150;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 2111 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 2052
Qy 772 CTGCTAATCCTGTTACCACTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
Db 2051 CTGCTAATCCTGTTACCACTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 1992
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGGCTGAAACGGGGGTTGTCGTGC 891
Db 1991 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGGCTGAAACGGGGGTTGTCGTGC 1932
Qy 892 ACACAGCCACCTTGGACGCAACCTACACCGAAGTACGATACCTACACGCGTGAGCAT 951
Db 1931 ACACAGCCACCTTGGACGCAACCTACACCGAAGTACGATACCTACACGCGTGAGCAT 1872
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 1871 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1812
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 1071
Db 1811 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 1752
Qy 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 1751 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1692
Qy 1132 CGGAGCCTATGAAAGAGCGGCAACGCGCGCGGG 1168
Db 1691 CGGAGCCTATGAAAGAGCGGCAACGCGCGCGGAG 1655

RESULT 13
A52326/c
LOCUS A52326 4410 bp DNA PAT 12-DEC-1997
DEFINITION Sequence 1 from Patent EP0725140.
ACCESSION A52326
VERSION A52326.1 GI:2851987

KEYWORDS unidentified.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4410)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Process of extraction of periplasmic proteins from prokaryotic
microorganisms in the presence of arginine
JOURNAL Patent: EP 0725140-A 1 07-AUG-1996;
COMMENT SANOFI SA (FR)
Other publication SK 10696 960904
Other publication Cz 9600290 960814
Other publication JP 8242879 960924
Other publication FI 960427 960801
Other publication PL 312543 960805
Other publication NO 960396 960801
Other publication FR 2729372 960802
Other publication CA 2168382 960801
Other publication AU 4224496 960808.
FEATURES
source Location/Qualifiers
1. .4410
/organism="unidentified"
/db_xref="taxon:32644"
promoter 1. .282
5'UTR 283. .337
misc_RNA 338. .762
terminator /note="SEQUENCE DE LA REGION 5' NON TRADUITE DU MESSAGE"
terminator /note="SEQUENCE CODANT POUR LE PRECURSEUR DE L'IL-13"
terminator 763. .812
misc_RNA 813. .1012
terminator /note="TERMINATEUR DU GENE 10 DU PHAGE T7"
terminator 1013. .1253
misc_RNA /note="TERMINATEUR DU PHAGE FD"
misc_RNA 1254. .2505
misc_RNA /note="GENE CODANT POUR LE REPRESSEUR DE L'OPERON LACTOSE"
misc_RNA 2506. .4410
BASE COUNT 1078 a 1142 c 1096 g 1094 t
ORIGIN

Query Match 31.8%; Score 453.8; DB 6; Length 4410;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 CTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 3023 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 2964
Qy 772 CTGCTAATCCTGTTACCACTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
Db 2963 CTGCTAATCCTGTTACCACTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 2904
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGGCTGAAACGGGGGTTGTCGTGC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGGCTGAAACGGGGGTTGTCGTGC 2844
Qy 892 ACACAGCCACCTTGGACGCAACCTACACCGAAGTACGATACCTACACGCGTGAGCAT 951
Db 2843 ACACAGCCACCTTGGACGCAACCTACACCGAAGTACGATACCTACACGCGTGAGCAT 2784
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2783 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2724
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 1071
Db 2723 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 2664
Qy 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 2663 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 2604

```
QY 1132 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGG 1168
|||||
Db 2603 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGAG 2567

RESULT 14
AR027062/c
LOCUS AR027062 4410 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5856142.
ACCESSION AR027062
VERSION AR027062.1 GI:5937902
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4410)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Method for the extraction of periplasmic proteins from prokaryotic
microorganisms in the presence of arginine
JOURNAL Patent: US 5856142-A 1 05-JAN-1999;
FEATURES
Location/Qualifiers
source 1..4410
BASE COUNT 1078 a 1142 c 1096 g 1094 t
ORIGIN

Query Match 31.8%; Score 453.8; DB 6; Length 4410;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 3023 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 2964

QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTCTTACCGGGTTG 831
Db 2963 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTCTTACCGGGTTG 2904

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGCTCGGGCTGAACGGGGGTTTCGTGC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGCTCGGGCTGAACGGGGGTTTCGTGC 2844

QY 892 ACACAGCCAGCTTGGAGCGAAGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 2843 ACACAGCCAGCTTGGAGCGAAGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 2784

QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2783 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2724

QY 1012 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGT 1071
Db 2723 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGT 2664

QY 1072 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 2663 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 2604

QY 1132 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGG 1168
|||||
Db 2603 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGAG 2567

Search completed: January 17, 2002, 11:49:01
Job time: 17153 sec
```

```
QY 1132 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGG 1168
|||||
Db 2603 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGAG 2567

RESULT 14
AR027062/c
LOCUS AR027062 4410 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5856142.
ACCESSION AR027062
VERSION AR027062.1 GI:5937902
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4410)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Method for the extraction of periplasmic proteins from prokaryotic
microorganisms in the presence of arginine
JOURNAL Patent: US 5856142-A 1 05-JAN-1999;
FEATURES
Location/Qualifiers
source 1..4410
BASE COUNT 1078 a 1142 c 1096 g 1094 t
ORIGIN

Query Match 31.8%; Score 453.8; DB 6; Length 4410;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 3023 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 2964

QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTCTTACCGGGTTG 831
Db 2963 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTCTTACCGGGTTG 2904

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGCTCGGGCTGAACGGGGGTTTCGTGC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGCTCGGGCTGAACGGGGGTTTCGTGC 2844

QY 892 ACACAGCCAGCTTGGAGCGAAGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 2843 ACACAGCCAGCTTGGAGCGAAGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 2784

QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2783 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2724

QY 1012 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGT 1071
Db 2723 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGT 2664

QY 1072 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 2663 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 2604

QY 1132 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGG 1168
|||||
Db 2603 CGGAGCCTATGGAAGAACGCCAGCAACCGCGCGGAG 2567

RESULT 15
I86195/c
LOCUS I86195 4410 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 1 from patent US 5700665.
ACCESSION I86195
VERSION I86195.1 GI:3205913
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:02:19 ; Search time 599.86 Seconds
(without alignments)
2036.623 Million cell updates/sec

Title: US-09-242-202a-16
Perfect score: 1425
Sequence: 1 TCCCATGCGGGGATCTTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_ll01:*

1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1425	100.0	1425	19	AAV21724 Humanised vector p
2	1203	84.4	2125	19	AAV21727 Humanised vector p
3	1194.4	83.8	1911	19	AAV21726 Humanised vector p
4	490.4	34.4	1547	19	AAV21732 Humanised vector p
5	459.8	32.3	3553	15	AAQ74208 Expression vector
6	454.8	31.9	7073	18	TF8-5G9 CDR-grafte
7	454.8	31.9	7864	18	AAV50963 TF8-5G9 CDR-grafte
8	454.8	31.9	11795	21	AAA59345 Nucleotide sequenc
9	454.8	31.9	13254	17	AAV40915 Nucleotide sequenc
10	454.8	31.9	13254	21	AAA32151 pEel2 Combo BM 12
11	454.8	31.9	13254	21	AAA32165 Complete nucleotid

c	12	453.8	31.8	608	19	AAV21731	ColEI origin of re
c	13	453.8	31.8	5241	17	AAV27307	pHL104 containing
c	14	453.8	31.8	5365	10	AAV90646	Nucleotide sequenc
c	15	453.8	31.8	5413	10	AAV90649	Nucleotide sequenc
c	16	453.8	31.8	5518	10	AAV90647	Nucleotide sequenc
c	17	453.8	31.8	5566	10	AAV90648	Nucleotide sequenc
c	18	453.8	31.8	5579	11	AAQ02032	plasmid pHS 164.
c	19	453.8	31.8	6051	12	AAQ14934	pBG391 sequence.
c	20	453.8	31.8	6149	10	AAV90645	Sequence of T4 lym
c	21	453.8	31.8	6151	10	AAV90644	Sequence of T4 lym
c	22	453.8	31.8	6151	11	AAQ03006	Full length T4 CDN
c	23	453.8	31.8	6151	11	AAQ05608	Plasmid pBG381 inc
c	24	453.8	31.8	6165	13	AAQ20324	Sol.rhesus CD4 seq
c	25	453.8	31.8	6477	11	AAQ02030	plasmid pinf 4-49.
c	26	453.8	31.8	7252	20	AAV86928	Complete sequence
c	27	453.8	31.8	7377	11	AAQ03005	Full length T4 CDN
c	28	453.8	31.8	7377	11	AAQ05607	Plasmid p170-2 inc
c	29	453.8	31.8	8119	17	AAV33477	Subtilisin N62b/G1
c	30	453.4	31.8	1547	19	AAV21732	Humanised vector p
c	31	453.4	31.8	1807	19	AAV21733	Humanised vector p
c	32	453.4	31.8	2308	19	AAV21734	Humanised vector p
c	33	453.4	31.8	7507	16	AAQ91272	Plasmid vector pad
c	34	453.2	31.8	13254	21	AAA31025	pEel2 Combo BM 12
c	35	453	31.8	2077	20	AAV23771	Vector PASlib DNA.
c	36	453	31.8	2462	21	AAA74638	Plasmid pSP72. Es
c	37	453	31.8	2577	14	AAQ36620	Expression plasmid
c	38	453	31.8	2704	20	AAQ06758	Vector pUC38 nucle
c	39	453	31.8	2927	17	AAV29158	Plasmid pTRP. Syn
c	40	453	31.8	2927	17	AAV31789	Plasmid pTRP. Syn
c	41	453	31.8	2939	16	AAQ87350	Plasmid BGINV. Sy
c	42	453	31.8	2939	19	AAV37293	pBGINV plasmid use
c	43	453	31.8	3003	11	AAQ05745	Plasmid P.L-mu-smc
c	44	453	31.8	3104	10	AAV90296	DNA target sequenc
c	45	453	31.8	3130	19	AAV46334	Universal reporter

ALIGNMENTS

RESULT	1
AAV21724	AAV21724 standard; cDNA; 1425 BP.
ID	AAV21724 standard; cDNA; 1425 BP.
XX	
AC	AAV21724;
XX	
XX	17-AUG-1998 (first entry)
DF	Humanised vector pITL.
DE	Vector; vaccine; tumour; antigen; plasmid pITL; ds.
XX	
XX	Chimeric - Homo sapiens.
XX	Chimeric - Escherichia coli.
XX	Synthetic.
PH	Key Location/Qualifiers
FT	misc_feature 1..221
FT	/*tag= a
FT	/note= "stuffer sequence"
FT	polyA_site 22..481
FT	/*tag= b
FT	/note= "combined splice and polyA sequences"
FT	CDS 495..701
FT	/*tag= c
FT	/note= "SupF gene"
FT	misc_feature 712..1164
FT	/*tag= d
FT	/note= "ColEI origin of replication"
FT	promoter 1177..1425
FT	/*tag= e
FT	/note= "RANTES promoter"
XX	
XX	WO9806863-A1.
PN	

XX 19-FEB-1998.
 PD
 XX 14-AUG-1997; 97WO-US14306.
 PF
 XX 14-AUG-1996; 96US-0023931.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Nelson EL, Nelson PJ;
 PI
 XX WPI; 1998-159552/14.
 DR
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 PT
 XX Claim 14; Page 29-30; 125pp; English.
 PS
 XX Plasmid pITL comprises a base vector for novel humanised
 CC polynucleotide vectors. Such vectors comprise a human-derived
 CC promoter or mammalian homologue which is functional in mammalian
 CC target tissue and cells and a sequence acceptance site (see
 CC AAV21735-36), which accepts cDNA products from RT-PCR cloning. They
 CC also contain minimal non-human components, such as a replication
 CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
 CC are necessary for production of the vector, as well as human-derived
 CC splice and polyA sequences (see AAV21723). The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX
 SQ Sequence 1425 BP; 325 A; 381 C; 404 G; 315 T; 0 other;

Query Match 100.0%; Score 1425; DB 19; Length 1425;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCATCGCGCGGATCTTCTTACTGATGATGAGTGGGACATATATGTTATCAGTGA 60
 DB 1 Tgcatacgcggtattcttactcactgataaagtgtggtgacataattattgtttatcagtga 60

QY 61 TAAAGTGTCAAGCATGACAAAGTTGACGCGAATACAGTGTCCGCGCGCTGAGCT 120
 DB 61 taaagtgtcaagcatgacaaagtgcagcgaataacagtgtatcgtgcgcgctggaact 120

QY 121 GTTGACGAGGTGCGGTAGACGGTGTGACGACACGCAAACTGGCGGAACGGTTGGGGGT 180
 DB 121 gttgaacgaggttcggttagacggtctgacgacacgcaaacgttcggaacggttgggggt 180

QY 181 GCAGACGCGCGCTTACTGGCACTTCAGACACAGCGGCGCCTTAAGGCGCATATGG 240
 DB 181 gcagcagcgcgcttacttgcacttcaggaacacagcgcgcccttaaggcgcatatgg 240

QY 241 TGAGTGGATGCCCTGACCGCGGGATGGGGAGACTGTAGTCAGAGCCCGCGGCG 300
 DB 241 tgagtggatgcttgaccacagcgggatggggagacctgtatcagagcccccgggc 300

QY 301 AGCACAGGCAATGCCGCTCTCCCTGCAGTAGTAGTACTGCCGGGTGGGATCCC 360
 DB 301 agcacaggccaatgccgctctctccctgcagtgcagtgcagtgcaggtgggatccc 360

QY 361 TGTGACCCCTCCCGAGTGCCTCTCTCGGCCCTGGAAGTTGCACCTCCAGTGCACACG 420
 DB 361 tgtgacctctcccgagtgctctctctgcccctggaggttgcacctccagtgcccaaccagc 420

QY 421 CTGTGCTTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTCATATAATTA 480

DB 421 ctgtgcttaataaaattaaagttgcatcttttgcagactaggtgtgctctataataatta 480
 QY 481 TAAGCTTGATATCGAATCTTCTTCAACGTAACACTTTTACAGGGGCGTCACTTTGATAT 540
 DB 481 taagcttgatatacgaattcttctcaacgtaaacactttacagcgcgctattgatat 540
 QY 541 GATGCGCGCGCTTCCGATAAAGGAGCAGCGCAGTAAGACATTTACCCGTGTGGGGTT 600
 DB 541 gatgcgccccgcttccgataaaggagcagggccagtaaaagcattaccctggtggggtt 600
 QY 601 CCGAGCGGCCAAAGGAGCAGACTCTTAAATCTGCCGTATCTGACTTTCGAAGTTCGAAT 660
 DB 601 ccgagcgcccaaggagcagacttaaatctgcgctcatcagacttcgaaggttcgaat 660
 QY 661 CTTTCCCGCCACACCATCTCTTCAAAAGTCCGAAAGAAATCTCTGACGCCGCTGTAGCCG 720
 DB 661 ctttccccaccaccatcttcaaaagtcgaaagaattctcgcagccggtgtagccg 720
 QY 721 TAGTTAGGCCACCATCTTCAAGAACTCTGTAGCACCGCCTACATACCTGCTCTGCTAATC 780
 DB 721 tagttaggccaccactcaagaactctgtagcaccgcctacatacctcgtctctgtaactc 780
 QY 781 CTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTTACCGGGTTGGACTCAAGA 840
 DB 781 ctgttaccagtgctgctgccagtgcgataagtcgtcttaccggttggactcaaga 840
 QY 841 CGATAGTTCACCGATTAAGCGGAGCGGCTGGGCTGAACGGGGGTTCTGTCACACAGCCC 900
 DB 841 cgatagttaccggataaaggcgagcggtcggtgaacgggggttcgtgcacagagccc 900
 QY 901 AGCTTGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCTGAGCATTTGAGAAAGC 960
 DB 901 agcttggagcaacgacctacaccgaactgagatacctacagctgagcatcgaaaagc 960
 QY 961 GCCACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCCGGTAAAGCGGAGGTCGGAACA 1020
 DB 961 gccacgcttcccgaaggagaaagcgagcaggtatccggttaagcgcgaggtcggaaca 1020
 QY 1021 GGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTCTGCGGG 1080
 DB 1021 ggagagcgcacgagggagagcttcccagggggaacgcctgatatcttagtctctcggg 1080
 QY 1081 TTTGCCCACTCTGACTTGTAGCGTCTGATTTTGTGATCTCTCAGGGGGCGGAGCCTA 1140
 DB 1081 tttcgccactctgacttgagcgtcgatcttggatgctgcagggggcggaagccta 1140
 QY 1141 TGGAAAACGCCAGCAACGCGCGCGGGGATCCGGAGAGCTCCTCTAGATGAGAGACA 1200
 DB 1141 tggaaaaacgccagcaacgcgcccggggatccggagagctcaactctagatgagagaca 1200
 QY 1201 GTGAGGAGAGACAGAGACTCGAATTTCCGGAGCTATTTCAGTTTCTTTCCTTTTGT 1260
 DB 1201 gtgagggagagacagagactcgaaattccggagactatttcagtttcttctccgtttgt 1260
 QY 1261 GCAATTTCACTTATGATACCGGCCCAATCTTGGTGTCTATTTTGAAACTCCCTTAGGG 1320
 DB 1261 gcaatttcacttatgataccgcccgaatgcttggtgtatttttgaaactccccctttaggg 1320
 QY 1321 GATGCCCTCAACTGGCCCTATTAAGGCCAGCCTGAGCTGCAGAGGATTCCTGCAGAG 1380
 DB 1321 gatgccctcaactggccctataaaggccagcctgagctgacagaggtattcctgcagagg 1380
 QY 1381 ATCAAGACAGCAGCTGGACCTCGCACAGCCTCTCCACAGGTACC 1425
 DB 1381 atcaagacagcacgtggacctcgcacagcctctctccacaggtacc 1425

RESULT 2
 AAV21727
 ID AAV21727 standard; cDNA; 2125 BP.
 XX
 AC AAV21727;

```

XX 17-AUG-1998 (first entry)
XX Humanised vector pITL-hHER/neu.
DE
XX
XX Vector; vaccine: tumour; antigen: plasmid pITL-hHER/neu;
KW human; HER-2/neu; C-erbB-2; breast cancer; ds.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 13..921
XX /*tag= a
XX /product= human HER-2/neu
XX polyA_site 922..1181
XX /*tag= b
XX /note= "combined splice and polyA sequences"
XX CDS 1195..1401
XX /*tag= c
XX /note= "SupF gene"
XX misc_feature 1412..1864
XX /*tag= d
XX /note= "Cole1 origin of replication"
XX promoter 1877..2125
XX /*tag= e
XX /note= "RANTES promoter"
XX
XX W09806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Example 7; Page 41-42; 125pp; English.
XX
XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
XX a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
XX evaluate the toxicity of anti-tumour vaccination in rats, and in
XX phase I and phase II trials to evaluate polynucleotide vaccination
XX in advanced breast cancer. Novel humanised vectors, which can be
XX based on pITL, comprise a human-derived promoter or mammalian
XX homologue which is functional in mammalian target tissue and cells
XX and a sequence acceptance site which accepts cDNA products from
XX RT-PCR cloning. The vectors are non-replicating in mammalian cells
XX but are capable of extended stable expression of the target
XX sequence, generating an immune response in immunised individuals.
XX The vectors selectively elicit immune responses to the target
XX sequences with little or no immune response to the other components
XX of the vectors.
XX
XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;
XX
XX
XX Query Match 84.4%; Score 1203; DB 19; Length 2125;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 223 GCCTTAGGGCCATATGGTGAAGTGGATGCGTTCAGCCCGGGGATGGGGGAGACCTG 282
XX }
XX 922 gccttaagggccatgagtgagtgagtccttgaccccccagcggggagacctg 981

```

```

QY 283 TAGTCAGAGCCCGCGGACAGCACAGGCAATGCGCGCTTCCTCCCTGCAGTGAGTAGTGA 342
DB tagtcagagcccccgggcagcacagggccaaTgcccgctcttccccctcagtgagtagtga 1041
QY 343 CTGCCCCGGGTGGATCCCTGTGACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 402
DB ctgccccgggtgggataccctgtgacccctccccctccccctccccctccccctggcc 1101
QY 403 ACTCCAGTGGCCACAGCCTTGTCTTAATAAAATTAAGTTGCATCATTTTGTCTGACTAG 462
DB actccagtggccaccagccttgtcttaataaaattaaagtgtcatcattttgtctgactag 1161
QY 463 GTGTCTCTATAATATTATTAAGCTTGATATGAATCTTTCTCAAGCTTAACACATTTACAG 522
DB gtgtctctataataattattataagcttgatataagcttgatataagcttgatataagcttgat 1221
QY 523 CGCGCGCTCATTTTGTATATGATGCGCCCGCTTCCCGATTAAGGAGCAGCGCCAGTAAAGC 582
DB cgcgcgctcatttgatgatgacgccccgcttcccgataagggagcagggccagtaaaagc 1281
QY 583 ATTACCCGTGTGGGTTCGCCAGCGGCCAAAGGAGCAGACTCTAATCTGCCGTATC 642
DB attacccggtgtggggttcccgagcggccaaaggagcagactctaaatctgcgctcatc 1341
QY 643 GACTTCGAAGGTTTGAATCTCTTCCCTCCACACCATCACTTTCAAAAGTCCGAAAGATTC 702
DB gacttcgaaggttcgaatctcttccccccaccactcatcttcaaaagtcctcgaagaaatc 1401
QY 703 CTGAGCCCGCTGTAGCCGTAGTTAGGCCACCATCTCAAGAACTCTGTAGCACCCCTACA 762
DB ctgagccccgtgtagccgttagttagccaccacttaagaactctgtagcaccgctaca 1461
QY 763 TACCTCGCTGTCTAATCTCTTACAGTGGCTGCTGCCAGTGGGATAGTCTGTGCTT 822
DB tacctcgctgtctaactctgttaccagtggtctgcadtggtgcgataaagtcgtgtctt 1521
QY 823 ACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTTCGGCTGACCGGG 882
DB accggttggactcaagacgatagttaccggataaaggcgcagcgttcggctgaaacgggg 1581
QY 883 GGTTCGTGCACACGCCAGCTTGGAGCGAACGACCTACACCGCACTGAGATACCTACAG 942
DB ggttcgtgcacacgccagcttggagcgaacgacctacacgcgaactgagatacctacag 1641
QY 943 CGTGAGCATTTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTA 1002
DB cgtgagcattgagaaagcgcacgcttcccgaaaggagaaaggcggacaggtatccggta 1701
QY 1003 AGCGGAGGGTCCGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGGTAT 1062
DB agcggaggggtcggaaacagagagcgcacagggagagcttccaggggggaaacgcctggtat 1761
QY 1063 CTTTATAGTCTCTCGGGTTTTCGCCACTCTGACTTGAAGCTGCGATTTTGTGATGCTCG 1122
DB ctttatagctctcggttctcgccacctctgactgagcgtcgtatcttctgtgctgcg 1821
QY 1123 TCAGGGGGGGAGCCCTATGAAAAACCCAGAACCGCGCGCGGGGATCCGGAGAGCTC 1182
DB tcagggggggagccctatggaaaaacgccagcaacgcggccgggggatccggagagctc 1881
QY 1183 ACTCTAGATGAGAGCAGTGTGAGGAGACAGACACTCGCAATTTCCGAGCTATTTTCAG 1242
DB actctagatgagagagcagtgagggagagacagagactcgaatttccggagctatttcag 1941
QY 1243 TTTTCTTTTCCGTTTGTGCAATTTTCACTTATGATACCGGCAATGCTTGTGCTATTT 1302
DB tttctttccggtttgtgcaatttcacttatgataccgcgccaatgcttgggtgctattt 2001
QY 1303 TGGAAACTCCCTTAGGGGATGCCCTCACTGACCTGACCTGACCTGACCTGACCTGAC 1362
DB tggaaactcccccttaggggatgccccctcaactgccccctcaactgccccctcaactgccccctgagctgc 2061

```


|||||
Db 1556 gtcctgtcgggttcgcacactctgaactgagcgtcgattttgtgatgcctgcagggg 1615
QY 1130 GCGGAGCCTATGAAACCCAGACACGCGCGGGGATCCGGAGAGCTACTCTAG 1189
Db 1616 ggcggagccctatgaaacccagcaacgcgcggggatccggagagctcactctag 1675
QY 1190 ATGAGAGCAGTGAAGGAGAGACAGACTCGAATTTCCGGAGCTATTTTCAGTTTCTT 1249
Db 1676 atgagagcagtgaggagagacagactcgaatttcggagcatttcagtttctt 1735
QY 1250 TTCGGTTTGTGCAATTTCACTTATGATACGGCCAAATGCTGTGCTATTTTGGAAAC 1309
Db 1736 ttcggtttgtgcaatttcacttatgataccgcgcaaatgcttggttctattttggaac 1795
QY 1310 TCCCTTAGGGGATGCCCTCAACTGCCCTATATAAGGGCCAGCTGAGCTGCAGAGGAT 1369
Db 1796 tccctctaggggatgcccctcaactggccctataaaggccagcctgagctgcagagat 1855
QY 1370 TCCTGCAGAGATCAAGACAGCAGCTGGACCTCGCAGAGCCTCTCCACAGGTACC 1425
Db 1856 tcctgcagagatcaagacagcagctgagacctgcacagcctctcccacaggtacc 1911

RESULT 4
AAV21732
ID AAV21732 standard; cDNA; 1547 BP.
AC AAV21732;
XX
XX 17-AUG-1998 (first entry)
XX Humanised vector pITL-A.
XX Vector; vaccine; tumour; antigen; plasmid pITL-A; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT CDS 472..680
FT /*tag= c
FT /note= "SupF gene"
FT misc_feature 686..1292
FT /*tag= d
FT /note= "ColE1 origin of replication"
FT promoter 1311..1547
FT /*tag= e
FT /note= "RANTES promoter"
XX
XX WO9806863-A1.
XX 19-FEB-1998.
XX 14-AUG-1997; 97WO-US14306.
XX 14-AUG-1996; 96US-0023931.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX Humanised polynucleotide vectors - comprising human derived promoter and sequence acceptance site, used for the production of vaccines
XX Claim 14; Page 56-57; 125pp; English.
XX Plasmid pITL-A comprises a base vector for novel humanised polynucleotide vectors. The vector in Escherichia coli DH 10-beta/p3 is deposited as ATCC 98401. Novel vectors comprise a

CC human-derived promoter or mammalian homologue which is functional
CC in mammalian target tissue and cells and a sequence acceptance site
CC (see AAV21735-36) which accepts cDNA products from RT-PCR cloning
CC They also contain minimal non-human components, such as a replication
CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
CC are necessary for production of the vector. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX
SQ Sequence 1547 BP; 344 A; 408 C; 425 G; 370 T; 0 other;

Query Match 34.4%; Score 490.4; DB 19; Length 1547;
Best Local Similarity 62.8%; Pred. No. 9.4e-126;
Matches 984; Conservative 0; Mismatches 401; Indels 183; Gaps 7;
QY 3 CCATGGCGCGGATTCTTTATACATGATAAGTTGGTGGACATATTATGTTTATCATGTGATA 62
Db 12 ccatggcgcggattctttatcactgataaagttggtggacattattgtttatcagtgata 71
QY 63 AAGTGTCAAGCATGACAAAAGTTGAGCGCAATACAGTATCCGTCCGCGCCCTGGACTGT 122
Db 72 aagttcaagcatgacaaagttgagccgaatacagtgatcgtgcgcggccctggaactgt 131
QY 123 TGAACGAGGTGCGGTGACAGCGTCTGACGACACGCAAACTGGCGAAGCTTTGGGGTGC 182
Db 132 tgaacgaggtgcgcgtgagacggtctgacgacacgcaactgctgagaaagctgtg9999tgc 191
QY 183 AGCAGCGCGCGCTTTACTGGCACTTCAGAAACAGCGCGCGCTTAAGGGCCATATGGTG 242
Db 192 agcagcgcgctttactgctcactcagaaacagcggcgccctaaaggccatagtggt 251
QY 243 AGTGGATGCCCTTGACCCCGAGCGGGGATGGGGAGACCTGTAGTCAGAGCCCGCGGCGAG 302
Db 252 agtggatgcttgaccccgagcgggagtggggagacgttagtcagagaccccgggcag 311
QY 303 CACAGGCCAATGCCGCTCTCCCTGCACTGAGTAGTACTGCCCGGTGGGATCCCTG 362
Db 312 cacagggccaatgcccgcgtctccctgcag----- 341
QY 363 TGACCCCTCCCGAGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTCCCGCAGCCT 422
Db 342 -gatgagtagtgagtgccctcctcctgcccctggaagtggccactccagtgccaccagcct 400
QY 423 TGTCTATAAATAAATTAGTTGCATCATTTTCTGCTGCTAGTGTCTCTATATAATTATA 482
Db 401 tgtctataaataaattagttgcatcattttgtgactaggtgctcctctataataata 460
QY 483 AGCTTGATATCGAATTTCTTCTCAACGTAACACTTTACAGCGCGCGCTCATTTGATATGA 542
Db 461 agcttgatatacgaaattcttcggacttttgaagtgatggtggtg999ga-aggattcga 519
QY 543 TGGCGCCCGCTTCCCGATAAGGAGCAGCGCCAGTAAAGAGCATTAACCGTGGTGGTTC 602
Db 520 accttcgaagtcgatgcgcgcagatttagagtcgtcctcttggcgcgtcgggaacccc 579
QY 603 CGAGCGGCCAAAGGAGCAGACTCTAAATCTGCCGTATCGACTTCGAGGTTTCAATC- 661
Db 580 accagggtaaatgtcttactggtcgtcctccttatcgggaagcgggagcgcacatcata 639
QY 662 CTTCCCGCCACCATCATCTTCAAAAGTCCGAAAGAAATTTCTTCGAGC-CCGTGTAGC 718
Db 640 aatgacgcgcgctgtaaaagtttacgtttgagaagaattcctgcagcccgccgctgtgc 699
QY 719 CGTAGTTAGGCCACCATCTCAAGAACTCTGTAG----- 752

Db 700 tggcgtttttccataggtctcgccccctgacgagcatcacaaaaatcgacgtcgaagtc 759
 QY 753 ACCGCTTACATACCTGCTCTGCTTAATCTGTTACAGTGGCTGCTGCCAGTGGCGATAA 812
 Db 760 agaggtggcgaaccgacagactataagataaccagcggtttcccccctggagctccc 819
 QY 813 GTCGTGCTTACCGGGTGGAGCTACAGAGATATACCGGATAAAGCGCGAGCGTGGG 872
 Db 820 tegtgcgtctctgttccgaccctgcgttaccggataccctgcgcctttctccctt 879
 QY 873 CTGACGGGGGTCTGTCACACAGCCAGCTTGGAGCGAAGCACCTACACCAACTGAG 932
 Db 880 cgggaagcgtggcgctttctcaatgtccagctgtaggtatctcagttcggttaggtcg 939
 QY 933 ATACCTACAGCTGAGCATGAGAAAGCGCACGCTTCC----- 971
 Db 940 ttcgctccaagctggcgtgtgtacagaaacccccctgtagcccgagcgtgcgccttat 999
 QY 972 -----CGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAG 1010
 Db 1000 ccggttaactatcgttctgagtcacaccccgtaagacacagacttatcgccactggcagcag 1059
 QY 1011 GTCGACAGAGAGCGGACAGGAGCTTCCAGGGGAAACCCCTGTTATATAG 1070
 Db 1060 ccactggttaacaggtatgacagcgcgaggtatgtagcgtgtctacagagttcttgaagt 1119
 QY 1071 TCCTGTCGGGTTTCGCCACCTGACTTGTAGCGTGCATTTTGTGTGTCGTCGACGGGG 1130
 Db 1120 ggtggcctaactacgctacactagaaggacagtatttggatctgcgtctctgctgaagc 1179
 QY 1131 CGGAGCGATGAAAA----- 1148
 Db 1180 cagttacccttcggaagaaggttggtagctcttgatcccggaacaaacaccacgcgtggtta 1239
 QY 1149 -----CGCCACACAGCGCGCGGGGAT 1171
 Db 1240 gcggtgtttttgttgcagcagcagattacgcgcagaaaaaaggatctggtgggat 1299
 QY 1172 CCGGAGAGCTCACTAGATGAGAGAGCAGTGGAGAGACAGAGACTCGAATTTCCGG 1231
 Db 1300 ccggaagctcactatagatagagacagtgagagagacagactcgaatttccgg 1359
 QY 1232 AGCTATTTCAGTTTCTTTTCCGTTTGTGCAATTTCACHTTATGATACCGGGCCAAATGCTT 1291
 Db 1360 agctattcagtttcttcttcggtttgtgcaatttcaacttatgatacccgccaaatgctt 1419
 QY 1292 GFTTGTATTTTGAACCTCCCTTAGGGGATGCCCTCAACTGGCCCTATATAAGGCCCA 1351
 Db 1420 ggttgcatttttggaactcccccttggggatgccccctcaactcgaactataaaggcca 1479
 QY 1352 GCCTGAGCTGCAGAGGATTCTCTGAGAGGATCAAGACAGCACGTGGACCTCGCACAGCCT 1411
 Db 1480 gctgagctgcagagattcctgcagaggtatcaagacagcagctggaacctcgcaagcct 1539
 QY 1412 CTCCCA 1419
 Db 1540 cteccaca 1547

RESULT 5
 AAQ/4208/c
 ID AAQ/4208 standard; cDNA; 3553 BP.
 XX
 AC AAQ/4208;
 XX
 DT 31-MAY-1995 (first entry)
 XX
 DE Expression vector VJneo.
 XX
 KW Influenza virus; antigen; specific; immune response; nucleoprotein;
 KW hemagglutinin; polymerase; matrix protein; non-structural protein;
 KW human; vaccine; ds.
 XX

OS Synthetic.
 XX WO9421797-A.
 XX 29-SEP-1994.
 XX 14-MAR-1994; 94WO-US02751.
 XX 18-MAR-1993; 93US-0032383.
 PR 08-JUL-1993; 93US-0089985.
 XX (MERI) MERCK & CO INC.
 PA (VICA-) VICAL INC.
 XX
 PI Donnelly JJ, Dworki VJ, Liu MA, Montgomery DL, Parker SE;
 PI Shiver JW, Ulmer JB;
 XX WPI; 1994-317017/39.
 DR
 XX Polynucleotide vaccine comprising influenza virus genes - for
 PT vaccination against more than one strain of influenza virus
 PT
 PS Claim 24; Page 105-07; 171pp; English.
 CC The sequences given in AAQ74206-08 represent expression vectors which
 CC were used for the expression of influenza virus genes. VJ contains
 CC the contains the ampr gene and VJneo contains the kanR gene in the
 CC same orientation. DNA constructs such as these, containing the
 CC influenza virus genes are capable of inducing the expression of an
 CC antigenic influenza virus gene product which induces a specific immune
 CC response upon introduction of the DNA construct into animal tissue in
 CC vivo and resultant uptake of the DNA construct by cells which express
 CC the encoded influenza gene. The encoded influenza virus gene encodes
 CC nucleoprotein, hemagglutinin, polymerase, matrix or non-structural
 CC human influenza virus gene products. The virus gene is operably linked
 CC to one or more control sequences for incorporation in a vaccine.
 XX
 SQ Sequence 3553 BP; 869 A; 911 C; 839 G; 934 T; 0 other;

Query Match 32.3%; Score 459.8; DB 15; Length 3553;
 Best Local Similarity 97.5%; Pred. No. 3.7e-117;
 Matches 467; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 686 AAAGTCCGAAAGAAATTCCTCGACGCCGCTGTAGCGCTAGTTAGGCCACCACTTCAAGAACT 745
 Db 2422 AGATACCAATACTGCTTCACGTAGTGTAGCGCTAGTTAGGCCACCACTTCAAGAACT 2363
 QY 746 CTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTTTACCAAGTGGCTGTGCCAGTG 805
 Db 2362 CTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTTTACCAAGTGGCTGTGCCAGTG 2303
 QY 806 GCGATAAGTCTGTCCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCCAGC 865
 Db 2302 GCGATAAGTCTGTCCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCCAGC 2243
 QY 866 GGTCCGGCTGAACGGGGGTTTGTGTGCACACAGCCAGCTTGGAGCGAAGACCTACACCG 925
 Db 2242 GGTCCGGCTGAACGGGGGTTTGTGTGCACACAGCCAGCTTGGAGCGAAGACCTACACCG 2183
 QY 926 AACTGAGATACCTACAGCTGAGCATTTGAGAAAGCCCGCTTCCCGAAGGAGGAAAGG 985
 Db 2182 AACTGAGATACCTACAGCTGAGCATTTGAGAAAGCCCGCTTCCCGAAGGAGGAAAGG 2123
 QY 986 CGGACAGGTATCCGGTAAAGCGGCGAGGTGCGAACACAGGAGCGCAGAGGAGCTTCCAG 1045
 Db 2122 CGGACAGGTATCCGGTAAAGCGGCGAGGTGCGAACACAGGAGCGCAGAGGAGCTTCCAG 2063
 QY 1046 GGGGAAACCGCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTC 1105
 Db 2062 GGGGAAACCGCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTC 2003
 QY 1106 GATTTTGTGATCTGCTCAGGGGGCGGAGCTATGGAACACGACGACGACGCGGCC 1164

QY 1072 CCGTGGGTTTCGCCACCTGACTTGGGCTCGATTTTGTGATGCTGCTCAGGGGG 1131
 Db 2969 CCGTGGGTTTCGCCACCTGACTTGGGCTCGATTTTGTGATGCTGCTCAGGGGG 2910
 QY 1132 CGGAGCCTATGAAACCGCCAGCAACCGCGCGGGG 1169
 Db 2909 CGGAGCCTATGAAACCGCCAGCAACCGCGCGGGG 2872

RESULT 7

AAT50963/c
 ID AAT50963 standard; DNA; 7864 BP.

XX AC AAT50963;

XX DT 01-APR-1997 (first entry)

XX DE TF8-5G9 CDR-grafted light chain gene in pEel2TF8LCDR3.

XX KW Tissue factor; complementarity determining region; CDR;
 humanised antibody; monoclonal antibody; TF8-5G9; coagulation;
 thrombosis; restenosis; arteriosclerosis; therapy;
 pEel2TF8LCDR3; vector; ss.

XX OS Chimeric - Mus sp.

XX OS Chimeric - Homo sapiens.

XX OS Chimeric - Escherichia coli.

XX OS Chimeric - Simian virus 40.

XX OS Chimeric - Human cytomegalovirus strain AD 169.

XX OS Chimeric - Cricetus cricetus.

XX FH Key Location/Qualifiers

FT misc_RNA

FT 1..759

FT /tag= a

FT /product= TF8LCDR3 CDR-grafted heavy chain

FT misc_signal

FT 6..8

FT /tag= b

FT /function= "Kozak sequence"

FT sig_peptide

FT 9..68

FT /tag= c

FT 69..392

FT /tag= d

FT /product= CDR-grafted variable region

FT 393..710

FT /tag= e

FT /product= human kappa constant region

FT 711..753

FT /tag= f

FT 760..3284

FT /tag= g

FT /note= "includes polyA signal, ColE1 replication origin and ampicillin resistance genes"

FT misc_RNA

FT 3285..5736

FT /tag= h

FT /note= "includes CHO glutamine synthetase cDNA under control of the SV40 early promoter"

FT transformed bacteria"

FT misc_difference 4324

FT /tag= i

FT /note= "base 4324 is given as n in the specification"

FT misc_RNA

FT 5737..7864

FT /tag= j

FT /note= "includes hCMV middle intermediate early promoter and pSP64 polylinker"

XX WO9640921-A1.

XX 19-DEC-1996.

XX 06-JUN-1996; 96WO-US09287.

PR 07-JUN-1995; 95US-0480120.
 XX (JOHJ) JOHNSON & JOHNSON.
 XX Joliffe LK, Pulito VL, Zivin RA;
 DR WPI: 1997-099935/09.
 DR P-PSDB; AAW10233.
 XX CDR-grafted antibody capable of inhibiting human tissue factor -
 PT used in the treatment of thrombotic disorders, e.g. intravascular
 PT coagulation, arterial restenosis etc.
 XX Claim 22; Page 84-89; 142pp; English.
 XX Vector pEel2TF8LCDR3 (AAT50963) comprises the TF8LCDR3 CDR-grafted
 CC light chain gene inserted into vector pEel2. It encodes a
 CC CDR-grafted light chain (AAW10233) in which CDRs (see also AAW10225-
 CC 27) from the mouse anti-tissue factor monoclonal antibody TF8-5G9
 CC are grafted into a human kappa C region. Humanised antibodies
 CC are obtained from the human kappa C region. Humanised antibodies
 CC are obtained by cotransfecting a host cell with vectors encoding the
 CC CDR-grafted light and heavy (see also AAW10232) chains. These
 CC antibodies are capable of inhibiting human tissue factor and can be
 CC used to treat e.g. intravascular coagulation, arterial restenosis
 CC and arteriosclerosis.
 XX Sequence 7864 BP; 2020 A; 1925 C; 1850 G; 2068 T; 1 other;

Query Match 31.9%; Score 454.8; DB 18; Length 7864;
 Best Local Similarity 99.6%; Pred. No. 1.2e-115;
 Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 771
 Db 1728 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 1669
 QY 772 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGAATAAGTCTGTCTTACC GG GTT G 831
 Db 1668 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGAATAAGTCTGTCTTACC GG GTT G 1609
 QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCGTGC 891
 Db 1608 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCGTGC 1549
 QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAAGTGTGATACCTACAGCGTGAGCAT 951
 Db 1548 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAAGTGTGATACCTACAGCGTGAGCAT 1489
 QY 952 TGAGAAAGCGCCACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 1011
 Db 1488 TGAGAAAGCGCCACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 1429
 QY 1012 GTCCGACAGAGAGCGCAGGAGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGT 1071
 Db 1428 GTCCGACAGAGAGCGCAGGAGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGT 1369
 QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGGCGTCCGATTTTGTGATGCTGCTCAGGGGG 1131
 Db 1368 CCTGTCGGGTTTCGCCACCTCTGACTTGGCGTCCGATTTTGTGATGCTGCTCAGGGGG 1309
 QY 1132 CGGAGCCTATGGAACCAACGCCAGCAACCGCGCGGGGG 1169
 Db 1308 CGGAGCCTATGGAACCAACGCCAGCAACCGCGCGGGGG 1271

RESULT 8

AAA59345/c

ID AAA59345 standard; DNA; 11795 BP.

XX AC

XX AAA59345;

XX

DT 07-NOV-2000 (first entry)
XX Nucleotide sequence of plasmid pNIV4801.
XX Fusion protein; Varicella Zoster Virus; VZV; gE protein; chicken pox;
XX shingles; NSI protein; ss.
XX Synthetic.
OS Varicella Zoster Virus.
XX WO200043527-A1.
XX 27-JUL-2000.
XX 17-JAN-2000; 2000WO-EP00352.
XX 20-JAN-1999; 99GB-0001254.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Bollen A, Haumont M, Jacobs P, Jacquet A, Massaer MGF;
XX WPT; 2000-505841/45.
XX Fusion protein comprising Varicella Zoster Virus gE protein and a
XX different VZV protein useful for treating Varicella (chickenpox) or
XX Zoster (shingles) -
XX Disclosure; Page 27-34; 60pp; English.
XX The specification describes a fusion protein, which comprises
XX a Varicella Zoster Virus (VZV) gE protein or an immunologically
XX active fragment, fused to a different protein of VZV. The
XX fusion protein or nucleic acids encoding it can be used to
XX prevent or ameliorate Varicella (chicken pox) or Zoster (shingles)
XX infections. The present sequence represents the plasmid pNIV4801, a
XX vector which is used to express fusion proteins of the invention.
XX Sequence 11795 BP; 2917 A; 2760 C; 2957 G; 3161 T; 0 other;

Query Match 31.9%; Score 454.8; DB 21; Length 11795;
Best Local Similarity 99.6%; Pred. No. 1.4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
DB 986 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 927
QY 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTACCAGGGTTG 831
DB 926 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTACCAGGGTTG 867
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGCTGAACGGGGGGTTCTGTGC 891
DB 866 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGCTGAACGGGGGGTTCTGTGC 807
QY 892 ACACAGCCACGTTGGAGCGAACGACCTACACCGAAGTACCTACACGCTGAGCAT 951
DB 806 ACACAGCCACGTTGGAGCGAACGACCTACACCGAAGTACCTACACGCTGAGCAT 747
QY 952 TGAGAAAGCGCCAGCGTTCCCGAAGGAGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 746 TGAGAAAGCGCCAGCGTTCCCGAAGGAGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG 687
QY 1012 GTCGGACAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 1071
DB 686 GTCGGACAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 627
QY 1072 CCTCTCGGGTTTCCCACTCTGACTGACGCTGATTTTTTCTGATGCTCTGCTCAGGGGGG 1131
DB 626 CCTCTCGGGTTTCCCACTCTGACTGACGCTGATTTTTTCTGATGCTCTGCTCAGGGGGG 567

QY 1132 CGGAGCCTATGAAAAACGCCAGCAACGCCGCGCGGGG 1169
DB 566 CGGAGCCTATGAAAAACGCCAGCAACGCCGCGCGGAGG 529
RESULT 9
AAT40915/c
ID AAT40915 standard; DNA; 13254 BP.
XX AC AAT40915;
XX 29-JAN-1997 (first entry)
DE Nucleotide sequence of pEel2 Combo BM 12.
XX Heavy chain; light chain; variable region; VH: monoclonal antibody;
KW MAb; HIV; human immunodeficiency virus; glycoprotein; gp120; clone;
KW virus infectivity assay; precursor gp160; immunocompetence; human;
KW anti-HIV antibody; detection; HIV infection; ds.
XX Synthetic.
OS
XX WO9602273-A1.
XX 01-FEB-1996.
XX 11-JUL-1995; 95WO-US08743.
XX 18-JUL-1994; 94US-0276852.
XX (SCRI) SCRIPPS RES INST.
XX Barbas CF, Burton DR, Lerner RA;
XX WPT; 1996-179601/18.
XX Monoclonal antibody binding to VL/V2 loop of HIV gp120 - used in
XX passive immuno:therapy and detection of HIV infection.
XX Example 3; Page 276-285; 366pp; English.
XX This sequence represents the nucleotide sequence of the pEel2 vector,
XX pEel2 Combo BM 12, which contains the b12 heavy and light chain genes.
XX The pEel2 vector has a human CMV promoter for expression of the light
XX chain, a polylinker to provide cloning sites, and a polyadenylation
XX signal for termination of transcription. The vector also contains the
XX GS selectable marker gene whose expression is controlled by an SV40
XX early promoter at the 5' end of the GS gene, an intron, and a
XX polyadenylation signal at the 3' end of the GS gene. A heavy chain
XX cassette comprising the HCMV promoter, enhancer elements, heavy chain
XX gene and polyadenylation signal were removed from the pEE6 vector and
XX inserted into the pEel2 vector to generate the combinatorial construct
XX containing both the b12 light and heavy chain genes. The vector pEel2
XX Combo BM 12 was used to transfect CHO cells and an antibody, b12, was
XX expressed. The resulting antibody has the capacity to reduce HIV
XX infectivity titre in an in vivo virus infectivity assay by 50 % at a
XX concentration of less than 700 ng of antibody/ml. The MAb may be used
XX for determining immunocompetence of a human anti-HIV antibody and in
XX the detection of HIV infection.
XX Sequence 13254 BP; 3206 A; 3559 C; 3251 G; 3238 T; 0 other;
Query Match 31.9%; Score 454.8; DB 17; Length 13254;
Best Local Similarity 99.6%; Pred. No. 1.5e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
DB 6449 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 6390
QY 772 GTGTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTACCAGGGTTG 831

Db 6389 CTGCTAATCCTGTTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 6330
QY 832 GACTCAAGACGATGTTACCGGATAAAGCGAGCGTGGCTGCGGTGAACGGGGTTCGTGC 891
Db 6329 GACTCAAGACGATGTTACCGGATAAAGCGAGCGTGGCTGCGGTGAACGGGGTTCGTGC 6270
QY 892 ACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCGCCAGCTTCCGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 1011
Db 6209 TGAGAAAGCGCCAGCTTCCGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 6150
QY 1012 CTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 1071
Db 6149 CTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 6090
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 6089 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 6030
QY 1132 CGGAGCCTATGGAAGAAAGCGGACCAACGCGCGGGGG 1169
Db 6029 CGGAGCCTATGGAAGAAAGCGGACCAACGCGCGGGGG 5992

RESULT 10

AAA32151/c
ID AAA32151 standard; DNA; 13254 BP.

XX
AC AAA32151;
XX

DT 04-JUL-2000 (first entry)

DE pBe12 Combo BM 12 containing b12 heavy and light chains.

XX Antibody; anti-HIV monoclonal antibody; glycoprotein-120;
KW human immunodeficiency virus type 1; HIV-1; infectivity titre;
KW passive immunotherapy; reduce severity; HIV-induced disease;
KW immunocompetence; active immunisation; ss.

OS Synthetic.

XX AU9948754-A.

XX 17-FEB-2000.

XX 16-SEP-1999; 99AU-0048754.

XX 16-SEP-1999; 99AU-0048754.

XX (SCRI) SCRIPPS RES INST.

XX Burton DR, Barbas CF, Lerner RA;

XX WPI; 2000-246867/22.

XX Human neutralizing monoclonal antibodies to human immunodeficiency
PT virus (HIV) used for providing passive immunotherapy to HIV are
PT specific for glycoprotein-120 -

PS Example 4; Figure 29; 374pp; English.

XX This sequence represents a polynucleotide used in the preparation of the
CC antibodies of the invention. The invention relates to the production of
CC an anti-HIV (human immunodeficiency virus) glycoprotein (gp)-120
CC monoclonal antibody capable of reducing an HIV infectivity titre in an
CC in vitro virus infectivity assay by 50% at a concentration of less than
CC 70 ng/mL. The method for the production of the antibody comprises:
CC (a) providing a first polynucleotide encoding a heavy chain
CC immunoglobulin amino acid sequence (which does not comprise the sequence
CC represented by AAY98206) and a second polynucleotide encoding a light

CC chain immunoglobulin amino acid sequence;
CC (b) inserting the first and second polynucleotide sequences into a host
CC cell;
CC (c) maintaining the host cell in conditions which allow the amino acid
CC sequences encoded by the polynucleotides to be expressed in the host
CC cell; and
CC (d) isolating the antibody comprising the heavy and light chain
CC immunoglobulin amino acid sequences from the host cell.
CC The anti-HIV gp-120 monoclonal antibody is used for providing passive
CC immunotherapy to HIV in a human. They can be administered to high-risk
CC patients to reduce the likelihood and/or severity of HIV-induced disease
CC and to patients who are already HIV-infected. The antibodies are used
CC for neutralising field isolates which provides information about the
CC immunocompetence of an immune response in HIV patients, for detecting
CC HIV in a biological fluid or tissue sample e.g. by radioimmunoassay, for
CC producing anti-idiotypic antibodies which can be used for active
CC immunisation and to screen human monoclonal antibodies to identify those
CC with the same binding specificity and to monitor the course of HIV
CC disease therapy by measuring the changes in concentration of HIV present
CC in the body or in body fluids by immunoassay. The anti-HIV gp-120
CC monoclonal antibodies are encoded by a human polynucleotide sequence and
CC when used in vivo for diagnosis and immunotherapy of HIV-induced disease
CC reduce the problems of significant host immune response to the
CC antibodies associated with monoclonal antibodies of xenogeneic or
CC chimeric derivation.

XX
SQ Sequence 13254 BP; 3206 A; 3561 C; 3249 G; 3237 T; 1 other;

Query Match 31.9%; Score 454.8; DB 21; Length 13254;
Best Local Similarity 99.6%; Pred. No. 1.5e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAGAACTCTGTAGCACCGCTACATACCTCGCT 771
Db 6449 GTGTAGCGGTAGTTAGGCCACCACTTCAGAACTCTGTAGCACCGCTACATACCTCGCT 6390

QY 772 CTGCTAATCCTGTTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 831

Db 6389 CTGCTAATCCTGTTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6330

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGGTGAACGGGGGTTCTGTC 891

Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGGTGAACGGGGGTTCTGTC 6270

QY 892 ACACAGCCAGCTTGCAGCGAACCGACCTACACGGAAGTACCTACAGCGTGAGCAT 951

Db 6269 ACACAGCCAGCTTGCAGCGAACCGACCTACACGGAAGTACCTACAGCGTGAGCAT 6210

QY 952 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGAGAGGTATCCGGTAAGCGGAGG 1011

Db 6209 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGAGAGGTATCCGGTAAGCGGAGG 6150

QY 1012 GTGCGAACAGGAGAGCGACGAGGGAGCTTCCAGGGGAAACGCCCTGATCTTTATAGT 1071

Db 6149 GTGCGAACAGGAGAGCGACGAGGGAGCTTCCAGGGGAAACGCCCTGATCTTTATAGT 6090

QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131

Db 6089 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 6030

QY 1132 CGGAGCCTATGGAAGAAAGCGGACCAACGCGGGGG 1169

Db 6029 CGGAGCCTATGGAAGAAAGCGGACCAACGCGGGGG 5992

RESULT 11

AAA32165
ID AAA32165 standard; DNA; 13254 BP.

XX
AC AAA32165;
XX

DT 04-JUL-2000 (first entry)

XX DE Complete nucleotide sequence of b12 IgG1 expression vector.

XX DE Antibody: anti-HIV monoclonal antibody; glycoprotein-120;
 KW human immunodeficiency virus type 1; HIV-1; Infectivity titre;
 KW passive immunotherapy; reduce severity; HIV-induced disease;
 KW immunocompetence; active immunisation; ss.

XX OS Synthetic.

XX XX AU9948754-A.

XX PN 17-FEB-2000.

XX PD 16-SEP-1999; 99AU-0048754.

XX PF 16-SEP-1999; 99AU-0048754.

XX PR 16-SEP-1999; 99AU-0048754.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Burton DR, Barbas CF, Lerner RA;

XX DR WPI; 2000-246867/22.

XX PT Human neutralizing monoclonal antibodies to human immunodeficiency

XX PT virus (HIV) used for providing passive immunotherapy to HIV are

XX PT specific for glycoprotein-120

XX PS Disclosure; Page 291-300; 374pp; English.

XX CC This sequence represents a polynucleotide used in the preparation of the

XX CC antibodies of the invention. The invention relates to the production of

XX CC an anti-HIV (human immunodeficiency virus) glycoprotein (gp)-120

XX CC monoclonal antibody capable of reducing an HIV infectivity titre in an

XX CC in vitro virus infectivity assay by 50% at a concentration of less than

XX CC 70 ng/ml. The method for the production of the antibody comprises:

XX CC (a) providing a first polynucleotide encoding a heavy chain

XX CC immunoglobulin amino acid sequence (which does not comprise the sequence

XX CC represented by AAY98206) and a second polynucleotide encoding a light

XX CC chain immunoglobulin amino acid sequence;

XX CC (b) inserting the first and second polynucleotide sequences into a host

XX CC cell;

XX CC (c) maintaining the host cell in conditions which allow the amino acid

XX CC sequences encoded by the polynucleotides to be expressed in the host

XX CC cell; and

XX CC (d) isolating the antibody comprising the heavy and light chain

XX CC immunoglobulin amino acid sequences from the host cell.

XX CC The anti-HIV gp-120 monoclonal antibody is used for providing passive

XX CC immunotherapy to HIV in a human. They can be administered to high-risk

XX CC patients to reduce the likelihood and/or severity of HIV-induced disease

XX CC and to patients who are already HIV-infected. The antibodies are used

XX CC for neutralising field isolates which provides information about the

XX CC immunocompetence of an immune response in HIV patients, for detecting

XX CC HIV in a biological fluid or tissue sample e.g. by radioimmunoassay, for

XX CC producing anti-idiotypic antibodies which can be used for active

XX CC immunisation and to screen human monoclonal antibodies to identify those

XX CC with the same binding specificity and to monitor the course of HIV

XX CC disease therapy by measuring the changes in concentration of HIV present

XX CC in the body or in body fluids by immunoassay. The anti-HIV gp-120

XX CC monoclonal antibodies are encoded by a human polynucleotide sequence and

XX CC when used in vivo for diagnosis and immunotherapy of HIV-induced disease

XX CC reduce the problems of significant host immune response to the

XX CC antibodies associated with monoclonal antibodies of xenogeneic or

XX CC chimeric derivation.

XX SQ Sequence 13254 BP; 3238 A; 3231 C; 3579 G; 3206 T; 0 other;

Query Match 31.9%; Score 454.8; DB 21; Length 13254;

Best Local Similarity 99.6%; Pred. No. 1.5e-115; Indels 0; Gaps 0;

Matches 456; Conservative 0; Mismatches 2;

712 GTGTAGCGGTAGTTAGGCGCACCTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 771

Db 6806 gtgtagcgttagttaggccaccacttcaagaactctgtagcaccgctacatactcgct 6865
 Qy 772 CTGCTAATCCTGTTTACCAGTGGCTGCTGCCAGTGGCGGAGTAAGTGTCTTTACCGGGTTG 831
 Db 6866 ctgtaatacctgttaccagtggtctgctgccagtgctgataagtgctgtcttaccgggttg 6925
 Qy 832 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGTTCGTGC 891
 Db 6926 gactcaagacgatatgttaccgataaaggcagcgtctgaaagggttctgtgc 6985
 Qy 892 ACACAGCCCGAGCTTGGAGCGCAACGCTTACACGAACTGAGATACCTACACGCTGAGCAT 951
 Db 6986 acacagcccgcttgagcgaacgacctacacgaaactagatacctacagcgtgagcat 7045
 Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
 Db 7046 tgagaaagcgcacgcttcccgaaaggagaaagcgacaggtatccggttaagcggcagg 7105
 Qy 1012 GTCGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
 Db 7106 gtcggaacagagagcgcagcagagagcttccaggggaaacgctgtatctttatagt 7165
 Qy 1072 CCTGTCGGGTTTCCGCACTCTGACTTGAGCGTCAATTTTGTGATGCTGCTCAGGGGGG 1131
 Db 7166 cctgtcgggttctgcacctctgacttgagcgtcgattttgtgtgctgctcagggggg 7225
 Qy 1132 CGGAGCCTATGGAAGAAACGCCAGCAACGCCGCGCGGGG 1169
 Db 7226 cggagcctatggaaaaacgccagcaacgcggcccgagg 7263

RESULT 12

AAV21731/c

ID AAV21731 standard; cDNA; 608 BP.

XX AAV21731;

DT 17-AUG-1998 (first entry)

DE Cole1 origin of replication from vector pBR327.

KW Vector; vaccine; Cole1; tumour; antigen; ds.

OS Escherichia coli.

PN W09806863-A1.

XX 19-FEB-1998.

PF 14-AUG-1997; 97WO-US14306.

XX 14-AUG-1996; 96US-0023931.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Nelson EL, Nelson PJ;

XX WPI; 1998-159552/14.

Humanised polynucleotide vectors - comprising human derived promoter and sequence acceptance site, used for the production of vaccines

PS Disclosure; Page 14; 125pp; English.

XX This DNA sequence comprises a minimal cole1 origin of replication

CC isolated from vector pVR327 by digestion with BstVI and AatI.

CC Novel humanised vectors of the invention (see AAV21724, AAV21727 and

CC AAV21732-34) comprise a human-derived promoter or mammalian homologue

CC which is functional in a mammalian target tissue and cells and an

CC acceptance site which accepts cDNA products from RT-PCR cloning.

CC They also contain minimal non-human components that are necessary

CC for production of the vector. The cole1 replication origin

CC provides plasmid replication and growth within permissive strains
 CC of Escherichia coli. The novel vectors are used to express target
 CC antigens, especially tumour antigens. They are non-replicating in
 CC mammalian cells but are capable of extended stable expression of
 CC target sequences generating an immune response in immunised
 CC individuals. The vectors selectively elicit immune responses to
 CC the target sequences with little or no immune response to the other
 CC components of the vectors. The target antigens are expressed as
 CC intracellular polypeptides or peptides and, as such, are processed
 CC as self polypeptides or peptides and appropriately presented on
 CC antigen presenting cells.
 XX
 SQ Sequence 608 BP; 130 A; 175 C; 160 G; 143 T; 0 other;

Query Match 31.8%; Score 453.8; DB 19; Length 608;
 Best Local Similarity 99.6%; Pred. No. 8.8e-116;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
 Db GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 398
 QY 772 CTGCTAATCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 831
 Db CTGCTAATCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 338
 QY 832 GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGCTTCGTGC 891
 Db GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGCTTCGTGC 278
 QY 892 ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 Db ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 218
 QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAACGGCGAGG 1011
 Db TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAACGGCGAGG 158
 QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTATAGT 1071
 Db GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTATAGT 98
 QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
 Db CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 38
 QY 1132 CGGAGCCTATGGAAGAAACGCGACAGCGGGCGGGG 1168
 Db CGGAGCCTATGGAAGAAACGCGACAGCGGGCGGGG 1

RESULT 13
 AAT27307
 ID AAT27307 standard; DNA; 5241 BP.
 XX
 AC AAT27307;
 XX
 DT 07-AUG-1996 (first entry)
 XX
 DE pH1104 containing mutated viral RNA.
 XX
 KW Vaccine; RNA virus; influenza virus; promoter; gene expression;
 KW attenuation; pH1104; pH1926; ds; cyclic.
 XX
 OS Synthetic.
 XX
 PN EP704533-A1.
 XX
 PD 03-APR-1996.
 XX
 PF 30-SEP-1994; 94EP-0115505.
 XX

PR 30-SEP-1994; 94EP-0115505.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Hobom G, Menke A, Neumann G;
 XX
 DR WPI; 1996-173036/18.
 XX
 PT Attenuated RNA virus with improved expression rate - useful as
 PT vaccine, e.g. against HIV, herpes virus, rhinovirus or
 PT cytomegalovirus
 XX
 PS Disclosure; Page 21-31; 38pp; English.
 XX

Plasmid pH1104 (AAT27307) is based on pH1926, which comprises hybrid
 CAT cDNA with flanking non-coding sequences from influenza viral RNA
 segments, inserted in antisense orientation between murine rDNA
 promoter and terminator sequences. The CAT reporter gene replaces
 a haemagglutinin coding sequence, retaining viral 5' and 3'
 untranslated sequences, which cooperatively constitute the vRNA
 promoter structure. vRNA 3' end mutations in pH1926 were
 created by PCR. Mouse B82 L cells were transfected with mutated
 constructs and subsequently with helper influenza A/FPV/Bratislava
 virus. A single nucleotide exchange (G3A) abolished promoter activity.
 A double mutant (G3A, U8C) resulted in significant CAT activity, while
 a triple mutant (G3A, C5U, U8C) further enhanced CAT activity. Thus
 certain specific modifications result in viruses with greatly enhanced
 expression rates. This may be utilized in viral vaccine vectors.
 XX
 SQ Sequence 5241 BP; 1264 A; 1320 C; 1379 G; 1278 T; 0 other;

Query Match 31.8%; Score 453.8; DB 17; Length 5241;
 Best Local Similarity 99.6%; Pred. No. 2e-115;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
 Db GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 1678
 QY 772 CTGCTAATCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 831
 Db CTGCTAATCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 1738
 QY 832 GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGCTTCGTGC 891
 Db GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGCTTCGTGC 1738
 QY 892 ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 Db ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 1858
 QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAAACGCTGTTATCTTTATAGT 1011
 Db TGAGAAAGCCACGCTTCCCGAAGGAGAAACGCTGTTATCTTTATAGT 1918
 QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTATAGT 1071
 Db GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTTATCTTTATAGT 1978
 QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
 Db CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 2038
 QY 1132 CGGAGCCTATGGAAGAAACGCGACAGCGGGCGGGG 1168
 Db CGGAGCCTATGGAAGAAACGCGACAGCGGGCGGGG 2075

RESULT 14
 AAN90646/C
 ID AAN90646 standard; DNA; 5365 BP.
 XX

AC AAN90645;
 XX 01-FEB-1991 (first entry)
 XX Nucleotide sequence of region encoding first 113 AA of plasmid pBG394
 DE including soluble T4-like (st4) polypeptide number 9 (st4#9).
 XX
 XX HIV; soluble T4-like polypeptide 9; immunotherapeutic; prophylactic;
 XX plasmid pBG394; diagnostic.
 XX Homo sapiens.
 XX W08901940-A.
 XX 09-MAR-1989.
 XX 01-SEP-1988; 88WO-US02940.
 XX 07-JAN-1988; 88US-0141649.
 XX (BIOJ) BIOGEN INC.
 XX Fisher RA, Gilbert W, Sato VL, Flavell RA, Maraganore JM, Liu TR;
 XX WPI; 1989-085519/11.
 XX
 XX DNA sequences coding for soluble T4-like polypeptide(s) -
 XX used in immuno:therapeutic and immunosuppressive compsns. and for
 XX preventing, treating or detecting AIDS
 XX Fig 19; ; 207pp; English.
 XX
 XX It is the nucleotide sequence of the plasmid pBG394. The sequence
 XX was isolated from 2 libraries: a lambda gt cDNA library derived from T
 XX cell tumour line REX and a lambda gt10 cDNA library derived from
 XX peripheral blood lymphocytes (PBL). For screening, a series of chemically
 XX synthesised antisense oligonucleotide DNA probes based on the known T4
 XX protein sequence was used. The polypeptide encoded is useful in
 XX immunotherapeutic, prophylactic and diagnostic compsns. It can be used
 XX to purify HIV from a sample.
 XX Sequence 5365 BP; 1300 A; 1380 C; 1355 G; 1330 T; 0 other;
 SQ
 Query Match 31.8%; Score 453.8; DB 10; Length 5365;
 Best Local Similarity 99.6%; Pred. No. 2e-115;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 712 GTGTAGCGGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCAGCGGCTGAAACGGGGGTTCTGTC 771
 Db 3975 GTGTAGCGGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCAGCGGCTGAAACGGGGGTTCTGTC 3916
 Qy 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 Db 3915 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 3856
 Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 891
 Db 3855 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 3796
 Qy 892 ACACAGCCCGAGCTTTGGAGCGAACACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 Db 3795 ACACAGCCCGAGCTTTGGAGCGAACACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 3736
 Qy 952 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
 Db 3735 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAAGCGGCAGG 3676
 Qy 1012 GTCGGAACAGAGAGCGCACAGGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1071
 Db 3675 GTCGGAACAGAGAGCGCACAGGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 3616
 Qy 1072 CCTGTGCGGGTTTCCGCCACCTCTGACTTGAGCGTCGATTTTGTGTGATGCTCTGTCAGGGGGG 1131

Db 3615 CCTGTGCGGGTTTCCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCTGTCAGGGGGG 3556
 Qy 1132 CGGAGCCTATGAGAAACGCCAGCAACGCCGCGGGG 1168
 Db 3555 CGGAGCCTATGAGAAACGCCAGCAACGCCGCGGGGAG 3519
 RESULT 15
 AAN90649/C
 ID AAN90649 standard; DNA; 5413 BP.
 XX AC AAN90649;
 XX 01-AUG-1990 (first entry)
 XX Nucleotide sequence of region encoding first 131 AA of plasmid pBG395
 DE including soluble T4-like (st4) polypeptide number 10 (st4#10).
 XX HIV; soluble T4-like polypeptide 10; immunotherapeutic; prophylactic;
 XX plasmid pBG395; diagnostic.
 XX Homo sapiens.
 XX W08901940-A.
 XX 09-MAR-1989.
 XX 01-SEP-1988; 88WO-US02940.
 XX 07-JAN-1988; 88US-0141649, US-094322.
 XX (BIOJ) BIOGEN INC.
 XX Fisher RA, Gilbert W, Sato VL, Flavell RA, Maraganore JM, Liu TR;
 XX WPI; 1989-085519/11.
 XX DNA sequences coding for soluble T4-like polypeptide(s) -
 XX used in immuno:therapeutic and immunosuppressive compsns. and for
 XX preventing, treating or detecting AIDS
 XX Disclosure; ; 207pp; English.
 XX
 XX It is the nucleotide sequence of the plasmid pBG395. The sequence
 XX was isolated from 2 libraries: a lambda gt cDNA library derived from T
 XX cell tumour line REX and a lambda gt10 cDNA library derived from
 XX peripheral blood lymphocytes (PBL). For screening, a series of chemically
 XX synthesised antisense oligonucleotide DNA probes based on the known T4
 XX protein sequence was used. The polypeptide encoded is useful in
 XX immunotherapeutic, prophylactic and diagnostic compsns. It can be used
 XX to purify HIV from a sample.
 XX Sequence 5413 BP; 1309 A; 1401 C; 1365 G; 1338 T; 0 other;
 SQ
 Query Match 31.8%; Score 453.8; DB 10; Length 5413;
 Best Local Similarity 99.6%; Pred. No. 2e-115;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 712 GTGTAGCGGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCAGCGGCTGAAACGGGGGTTCTGTC 771
 Db 4023 GTGTAGCGGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCAGCGGCTGAAACGGGGGTTCTGTC 3964
 Qy 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 Db 3963 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 3904
 Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 891
 Db 3903 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 3844
 Qy 892 ACACAGCCCGAGCTTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951

```
|||||
Db 3843 ACACAGCCAGCTTGGAGCGAAGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 3784
QY 952 TGAGAAAGCGCCACACGTTCCCGAAGGGAGAAAGCGGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 3783 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGGACAGGTATCCGGTAAGCGGCAGG 3724
QY 1012 GTCGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 3723 GTCGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 3664
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 3663 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 3604
QY 1132 CGGAGCCTATGGAAAAACGCCAGCAACGCCGCCGGG 1168
Db 3603 CGGAGCCTATGGAAAAACGCCAGCAACGCCGCCGGGAG 3567
```

Search completed: January 17, 2002, 12:03:06
Job time: 17773 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:52:06 ; Search time 221.34 Seconds
(without alignments)
1458.079 Million cell updates/sec

Title: US-09-242-202A-16
Perfect score: 1425
Sequence: 1 TGCCATGGCGCGGATCTTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	454.8	31.9	13254	1	US-08-276-852-156
C 2	454.8	31.9	13254	1	US-08-276-852-170
C 3	454.8	31.9	13254	1	US-08-899-575-156
C 4	454.8	31.9	13254	1	US-08-899-575-170
C 5	454.8	31.9	13254	1	US-08-899-575-156
C 6	454.8	31.9	13254	1	US-08-899-575-170
C 7	454.8	31.9	13254	5	PCT-US95-08743-156
C 8	454.8	31.9	13254	5	PCT-US95-08743-170
C 9	453.8	31.8	1905	1	US-08-594-469-9
C 10	453.8	31.8	1905	2	US-08-906-957-9
C 11	453.8	31.8	4410	1	US-08-594-469-1
C 12	453.8	31.8	4410	2	US-08-906-957-1
C 13	453.8	31.8	6151	5	PCT-US91-02954-12
C 14	453.8	31.8	8119	1	US-08-460-343B-1
C 15	453.8	31.8	8119	1	US-08-398-028B-1
C 16	453.8	31.8	8119	2	US-08-504-265B-1
C 17	453	31.8	2927	2	US-08-941-647A-1
C 18	453	31.8	2939	1	US-08-119-512-2
C 19	453	31.8	2939	3	US-08-488-015B-2
C 20	453	31.8	2939	3	US-08-814-412-17
C 21	453	31.8	3003	6	5182260-18
C 22	453	31.8	3104	1	US-07-415-307A-1
C 23	453	31.8	3104	1	US-08-371-320-1
C 24	453	31.8	3130	4	US-09-038-141-1
C 25	453	31.8	3249	1	US-08-507-455-4
C 26	453	31.8	3301	2	US-08-447-430A-42
C 27	453	31.8	3400	1	US-08-507-455-3

28	453	31.8	3423	2	US-08-447-430A-40	Sequence 40, Appl
29	453	31.8	3474	2	US-08-447-430A-41	Sequence 41, Appl
30	453	31.8	3474	2	US-08-318-837-10	Sequence 10, Appl
31	453	31.8	3585	1	US-08-362-670B-9	Sequence 9, Appl
32	453	31.8	3585	3	US-08-333-576C-9	Sequence 9, Appl
33	453	31.8	3585	4	US-08-808-324-9	Sequence 9, Appl
34	453	31.8	3585	5	PCT-US94-14030A-9	Sequence 9, Appl
35	453	31.8	3623	2	US-07-989-847-13	Sequence 13, Appl
36	453	31.8	3623	4	US-08-469-411-13	Sequence 13, Appl
37	453	31.8	3632	1	US-07-745-382-13	Sequence 13, Appl
38	453	31.8	3632	1	US-07-921-848-13	Sequence 13, Appl
39	453	31.8	3632	1	US-08-115-680-3	Sequence 3, Appl
40	453	31.8	3632	1	US-07-941-372-3	Sequence 3, Appl
41	453	31.8	3632	1	US-08-165-301A-13	Sequence 13, Appl
42	453	31.8	3632	3	US-08-810-436-13	Sequence 13, Appl
43	453	31.8	3632	5	PCT-US93-08247-3	Sequence 3, Appl
44	453	31.8	3632	5	PCT-US94-14179-13	Sequence 13, Appl
45	453	31.8	3656	1	US-08-232-463-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-08-276-852-156/c

; Sequence 156, Application US/08276852

; Patent No. 5652138

; GENERAL INFORMATION:

; APPLICANT: Burton, Dennis R

; APPLICANT: Barbas, Carlos F

; APPLICANT: Lerner, Richard A

; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS

; NUMBER OF SEQUENCES: 170

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Scripps Research Institute, Office of

; ADDRESSEE: Patent Counsel

; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,

; STREET: Mail Drop TPC8

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/276.852

; FILING DATE: 18-JUL-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/178,302

; FILING DATE: 30-SEP-1993

; APPLICATION NUMBER: US 07/954,148

; FILING DATE: 30-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

REGISTRATION NUMBER:	34,163
REFERENCE/DOCKET NUMBER:	SCR1452P
TELEPHONE:	619-554-2937
TELEFAX:	619-554-6312
INFORMATION FOR SEQ ID NO:	156:
SEQUENCE CHARACTERISTICS:	
LENGTH:	13254 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	double
TOPOLOGY:	circular
MOLECULE TYPE:	DNA (genomic)
US-08-276-852-156	

```

Query Match      31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTACCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 771
Db 6449 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 6390
QY 772 CTGCTAATCTCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
Db 6389 CTGCTAATCTCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6330
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGACGCGTCCGGGTGAACGGGGGTTCTGTCG 891
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGGACGCGTCCGGGTGAACGGGGGTTCTGTCG 6270
QY 892 ACACAGCCCGCTTGGAGCGACGCTACACCGCACTGAGATACCTACACGGTGAGCAT 951
Db 6269 ACACAGCCCGCTTGGAGCGACGCTACACCGCACTGAGATACCTACACGGTGAGCAT 6210
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAAAGCGGACAGTATCCGCTAAGCGGCAAG 1011
Db 6209 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAAAGCGGACAGTATCCGCTAAGCGGCAAG 6150
QY 1012 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGAAACGCGCTGCTATCTTTATAGT 1071
Db 6149 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGAAACGCGCTGCTATCTTTATAGT 6090
QY 1072 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGGCTGATTTTGTGATGCTCTGTCAGGGGG 1131
Db 6089 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGGCTGATTTTGTGATGCTCTGTCAGGGGG 6030
QY 1132 CGGAGCCTATGAAAGCGGACGCAACGCGCCCGGGG 1169
Db 6029 CGGAGCCTATGAAAGCGGACGCAACGCGCCCGGGG 5992

RESULT 2
US-08-276-852-170
; Sequence 170, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148

```

```

; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-170

Query Match      31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTACCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 771
Db 6806 GTGTACCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 6865
QY 772 CTGCTAATCTCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
Db 6866 CTGCTAATCTCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6925
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGACGCGTCCGGGTGAACGGGGGTTCTGTCG 891
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGGACGCGTCCGGGTGAACGGGGGTTCTGTCG 6985
QY 892 ACACAGCCCGCTTGGAGCGACGCTACACCGCACTGAGATACCTACACGGTGAGCAT 951
Db 6986 ACACAGCCCGCTTGGAGCGACGCTACACCGCACTGAGATACCTACACGGTGAGCAT 7045
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAAAGCGGACAGTATCCGCTAAGCGGCAAG 1011
Db 7046 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAAAGCGGACAGTATCCGCTAAGCGGCAAG 7105
QY 1012 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGAAACGCGCTGCTATCTTTATAGT 1071
Db 7106 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGAAACGCGCTGCTATCTTTATAGT 7165
QY 1072 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGGCTGATTTTGTGATGCTCTGTCAGGGGG 1131
Db 7166 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGGCTGATTTTGTGATGCTCTGTCAGGGGG 7225
QY 1132 CGGAGCCTATGAAAGCGGACGCAACGCGCCCGGGG 1169
Db 7226 CGGAGCCTATGAAAGCGGACGCAACGCGCCCGGGG 7263

RESULT 3
US-08-899-575-156/c
; Sequence 156, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA

```

```

; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-899-575-156

```

```

Query Match 31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 6449 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 6390

QY 772 CTGCTAATCTGTTACCAAGTGGCTGCGCAAGTGGCGGTAAGTCTCTTACCGGGTTG 831
Db 6389 CTGCTAATCTGTTACCAAGTGGCTGCGCAAGTGGCGGTAAGTCTCTTACCGGGTTG 6330

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGCTGCGGCTGAACGGGGGTTCTGTC 891
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGCTGCGGCTGAACGGGGGTTCTGTC 6270

QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGCGAAGCACTTACCGAACTGAGATACCTACAGCGTGAGCAT 6210

QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAGAAAGCGGACAGGATATCCGGTAAGCGGCAGG 1011
Db 6209 TGAGAAAGCCACGCTTCCCGAAGGAGAGAAAGCGGACAGGATATCCGGTAAGCGGCAGG 6150

QY 1012 GTCGGAAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGATCTTTATAGT 1071
Db 6149 GTCGGAAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGATCTTTATAGT 6090

QY 1072 CTTGTCGGGTTTCCGCACTCTGACTTGACGCTGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 6089 CTTGTCGGGTTTCCGCACTCTGACTTGACGCTGATTTTGTGATGCTCGTCAGGGGGG 6030

QY 1132 CGGAGCCTATGGAAGAAACGCGACAGCGCGCGCGGGG 1169
Db 6029 CGGAGCCTATGGAAGAAACGCGACAGCGCGCGCGGGG 5992

```

```

RESULT 4
US-08-899-575-170
; Sequence 170, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-899-575-170

```

```

Query Match 31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 6806 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 6865

QY 772 CTGCTAATCTGTTACCAAGTGGCTGCGCAAGTGGCGGTAAGTCTCTTACCGGGTTG 831
Db 6866 CTGCTAATCTGTTACCAAGTGGCTGCGCAAGTGGCGGTAAGTCTCTTACCGGGTTG 6925

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGCTGAACGGGGGTTCTGTC 891
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGCTGAACGGGGGTTCTGTC 6985

QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6986 ACACAGCCAGCTTGGAGCGAAGCACTTACCGAACTGAGATACCTACAGCGTGAGCAT 7045

```

QY 952 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 1011
Db 7046 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 7105
QY 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 7106 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 7165
QY 1072 CTTGTCGGGTTTCGCGCACCTCTGACTTGAGCGTCAATTTTGTGATGCTCTCAGGGGG 1131
Db 7166 CTTGTCGGGTTTCGCGCACCTCTGACTTGAGCGTCAATTTTGTGATGCTCTCAGGGGG 7225
QY 1132 CGGAGCCTATGAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1169
Db 7226 CGGAGCCTATGAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7263

RESULT 5

US-08-899-575-156/C
; Sequence 156, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-156

Query Match 31.9%; Score 454.8; DB 1; Length 13254;

Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTGTGTAGCAGCGCTACATACCTCGCT 771
Db 6449 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTGTGTAGCAGCGCTACATACCTCGCT 6390
QY 772 CTGCTAATCTCTTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
Db 6389 CTGCTAATCTCTTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 6330
QY 832 GACTCAAGCAGTAGTTACCGGATAAAGCGGACGCGTTCGGGCTGAACGGGGGTTTCTGTGC 891
Db 6329 GACTCAAGCAGTAGTTACCGGATAAAGCGGACGCGTTCGGGCTGAACGGGGGTTTCTGTGC 6270
QY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 1011
Db 6209 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 6150
QY 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 6149 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 6090
QY 1072 CTTGTCGGGTTTCGCGCACCTCTGACTTGAGCGTCAATTTTGTGATGCTCTCAGGGGG 1131
Db 6089 CTTGTCGGGTTTCGCGCACCTCTGACTTGAGCGTCAATTTTGTGATGCTCTCAGGGGG 6030
QY 1132 CGGAGCCTATGAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1169
Db 6029 CGGAGCCTATGAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5992

RESULT 6

US-08-899-575-170
; Sequence 170, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992

ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 6806 GTGTAGCGGTAGTTAGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 6865
Qy 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
Db 6866 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 6925
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTCTGTGC 891
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTCTGTGC 6985
Qy 892 ACACAGCCAGCTTGGAGGAGCAACCTACACCACTGAGATACCTACACGCTGAGCAT 951
Db 6986 ACACAGCCAGCTTGGAGGAGCAACCTACACCACTGAGATACCTACACGCTGAGCAT 7045
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 7046 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 7105
Qy 1012 GTCGGAACAGAGAGCGGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 1071
Db 7106 GTCGGAACAGAGAGCGGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 7165
Qy 1072 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGATTTTGTGATGCTCTGTCAGGGGG 1131
Db 7166 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGATTTTGTGATGCTCTGTCAGGGGGG 7225
Qy 1132 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGGG 1169
Db 7226 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGAGG 7263

RESULT 7
PCT-US95-08743-156/c
Sequence 156, Application PC/TUS9508743
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
PCT-US95-08743-156

Query Match 31.9%; Score 454.8; DB 5; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 6449 GTGTAGCGGTAGTTAGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 6390
Qy 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
Db 6389 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 6330
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTCTGTGC 891
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTCTGTGC 6270
Qy 892 ACACAGCCAGCTTGGAGGAGCAACCTACACCACTGAGATACCTACACGCTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGGAGCAACCTACACCACTGAGATACCTACACGCTGAGCAT 6210
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 6209 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 6150
Qy 1012 GTCGGAACAGAGAGCGGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 1071
Db 6149 GTCGGAACAGAGAGCGGACGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 6090
Qy 1072 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGATTTTGTGATGCTCTGTCAGGGGG 1131
Db 6089 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGATTTTGTGATGCTCTGTCAGGGGGG 6030
Qy 1132 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGGG 1169
Db 6029 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGAGG 5992

RESULT 8
PCT-US95-08743-170
Sequence 170, Application PC/TUS9508743
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

PCT-US95-08743-170

```

Query Match          31.9%; Score 454.8; DB 5; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCCTAGTTAGGCGCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
Db 6806 GTGTAGCCCTAGTTAGGCGCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 6865
QY 772 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 831
Db 6866 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 6925
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGTTCGGCTGAACGGGGGTTCTGTC 891
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGTTCGGCTGAACGGGGGTTCTGTC 6985
QY 892 ACACAGCCCGAGCTTGGCGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6986 ACACAGCCCGAGCTTGGCGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 7045
QY 952 TCAGAAAGCGCCAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
Db 7046 TCAGAAAGCGCCAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 7105
QY 1012 GTCGGAACAGAGAGCGGACAGGAGGAGCTTCCAGGGGGAACGCTGCTATCTTTATAGT 1071
Db 7106 GTCGGAACAGAGAGCGGACAGGAGGAGCTTCCAGGGGGAACGCTGCTATCTTTATAGT 7165
QY 1072 CCTGTCGGGTTTCCCGACCTCTGACTTGAGCGTGAATTTTGTGATGCTCGTACAGGGGG 1131
Db 7166 CCTGTCGGGTTTCCCGACCTCTGACTTGAGCGTGAATTTTGTGATGCTCGTACAGGGGG 7225
QY 1132 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGGGG 1169
Db 7226 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGGGG 7263

```

```

RESULT 9
US-08-594-469-9/c
; Sequence 9, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,469
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E

```

```

; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-594-469-9

Query Match          31.8%; Score 453.8; DB 1; Length 1905;
Best Local Similarity 99.6%; Pred. No. 6.3e-133;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCCTAGTTAGGCGCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
Db 518 GTGTAGCCCTAGTTAGGCGCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 459
QY 772 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 831
Db 458 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 399
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGTTCGGCTGAACGGGGGTTCTGTC 891
Db 398 GACTCAAGACGATAGTTACCGGATAAGCGGCGAGCGGTTCGGCTGAACGGGGGTTCTGTC 339
QY 892 ACACAGCCCGAGCTTGGAGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 338 ACACAGCCCGAGCTTGGAGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 279
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
Db 278 TGAGAAAGCGCCAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 219
QY 1012 GTCGGAACAGAGAGCGCAGCAGGAGGAGCTTCCAGGGGGAACGCTGCTATCTTTATAGT 1071
Db 218 GTCGGAACAGAGAGCGCAGCAGGAGGAGCTTCCAGGGGGAACGCTGCTATCTTTATAGT 159
QY 1072 CCTGTCGGGTTTCCCGACCTCTGACTTGAGCGTGAATTTTGTGATGCTCGTACAGGGGG 1131
Db 158 CCTGTCGGGTTTCCCGACCTCTGACTTGAGCGTGAATTTTGTGATGCTCGTACAGGGGG 99
QY 1132 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGGGG 1168
Db 98 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGGGG 62

```

```

RESULT 10
US-08-906-957-9/c
; Sequence 9, Application US/08906957
; Patent No. 5856142
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,957
; FILING DATE: 06-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/594,469
; FILING DATE:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-906-957-9

Query Match      31.8%; Score 453.8; DB 2; Length 1905;
Best Local Similarity 99.6%; Pred. No. 6.3e-133;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 712 GTGTAGCGGTAGTTAGCGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 518 GTGTAGCGGTAGTTAGCGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 459
Qy 772 CTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTTACCGGGTTG 831
Db 458 CTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTTACCGGGTTG 399
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTGCGGCTGAACGGGGGTTCTGTC 891
Db 398 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTGCGGCTGAACGGGGGTTCTGTC 339
Qy 892 ACACAGCCCGAGCTTGGAGCGGACCACTACACCGAAGTGTAGATACCTACAGCGTGAGCAT 951
Db 338 ACACAGCCCGAGCTTGGAGCGGACCACTACACCGAAGTGTAGATACCTACAGCGTGAGCAT 279
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 1011
Db 278 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 219
Qy 1012 GTCGGAACAGAGAGCGACGAGGAGCTTCCAGGGGAAACGCTTGTATCTTTATAGT 1071
Db 218 GTCGGAACAGAGAGCGACGAGGAGCTTCCAGGGGAAACGCTTGTATCTTTATAGT 159
Qy 1072 CCTGTCGGGTTTCCCGACCTCTGACTTTGAGCGTGCATTTTGTGATCGTCTGTCAGGGGG 1131
Db 158 CCTGTCGGGTTTCCCGACCTCTGACTTTGAGCGTGCATTTTGTGATCGTCTGTCAGGGGG 99
Qy 1132 CGGAGCCTATGGAAAGAAAGCGCAGCAACGCGCGCGGG 1168
Db 98 CGGAGCCTATGGAAAGAAAGCGCAGCAACGCGCGCGCGAG 62

RESULT 11
US-08-594-469-1/c
; Sequence 1, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul

```

```

; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; periplasmic proteins of prokaryotic microorganisms in the
; presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,469
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-594-469-1

```

```

Query Match      31.8%; Score 453.8; DB 1; Length 4410;
Best Local Similarity 99.6%; Pred. No. 1e-132;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 712 GTGTAGCGGTAGTTAGCGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 3023 GTGTAGCGGTAGTTAGCGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 2964
Qy 772 CTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTTACCGGGTTG 831
Db 2963 CTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTCTTTACCGGGTTG 2904
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTGCGGCTGAACGGGGGTTCTGTC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTGCGGCTGAACGGGGGTTCTGTC 2844
Qy 892 ACACAGCCCGAGCTTGGAGCGGACCACTACACCGAAGTGTAGATACCTACAGCGTGAGCAT 951
Db 2843 ACACAGCCCGAGCTTGGAGCGGACCACTACACCGAAGTGTAGATACCTACAGCGTGAGCAT 2784
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 1011
Db 2783 TGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 2724
Qy 1012 GTCGGAACAGAGAGCGACGAGGAGCTTCCAGGGGAAACGCTTGTATCTTTATAGT 1071
Db 2723 GTCGGAACAGAGAGCGACGAGGAGCTTCCAGGGGAAACGCTTGTATCTTTATAGT 2664
Qy 1072 CCTGTCGGGTTTCCCGACCTCTGACTTTGAGCGTGCATTTTGTGATCGTCTGTCAGGGGG 1131
Db 2663 CCTGTCGGGTTTCCCGACCTCTGACTTTGAGCGTGCATTTTGTGATCGTCTGTCAGGGGG 2604

```

QY 1132 CGAGCCTATGGAACACCCAGCAACGCGCGCGGG 1168
|||||
Db 2603 CGAGCCTATGGAACACCCAGCAACGCGCGCGAG 2567

RESULT 12

US-08-906-957-1/c
; Sequence 1, Application US/08906957
; Patent No. 5856142
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALONE, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Staters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,957
; FILING DATE: 06-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/594,469
; FILING DATE:
; FILING DATE: 31-JAN-1995
; APPLICATION NUMBER: FR 95 01083
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-906-957-1

Query Match 31.8%; Score 453.8; DB 2; Length 4410;
Best Local Similarity 99.6%; Pred. No. 1e-132;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTACCGCACTTCAAGAACTGTAGCAGCGGCTGAACGCGGCTTACATACCTCGCT 771
Db 3023 GTGTAGCCGTAGTACCGCACTTCAAGAACTGTAGCAGCGGCTGAACGCGGCTTACATACCTCGCT 2964
QY 772 CTGCTAATCTGTTACAGTGGCTGCTGCAGTGGGATAAGTCTGTCTTACCGGGTTG 831
Db 2963 CTGCTAATCTGTTACAGTGGCTGCTGCAGTGGGATAAGTCTGTCTTACCGGGTTG 2904
QY 832 GACTCAAGCAGTAGTACCGGATAAGCGGCTGCGGCTGAACGCGGCTTACATACCTCGCT 891
Db 2903 GACTCAAGCAGTAGTACCGGATAAGCGGCTGCGGCTGAACGCGGCTTACATACCTCGCT 2844
QY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAAGCAT 951
|||||

Db 2843 ACACAGCCAGCTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCTGAGCAT 2784
QY 952 TGAGAAAGCGCCAGCGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGAAGCGGAGG 1011
|||||
Db 2783 TGAGAAAGCGCCAGCGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGAAGCGGAGG 2724
QY 1012 GTCGACAGGAGAGCGGACAGGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 1071
|||||
Db 2723 GTCGAAACAGGAGAGCGGACAGGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 2664
QY 1072 CCTGTGCGGTTTCCCGACCTCTGACTTGAAGTGGCTGATTTTGTGATGCTGTCAGGGGG 1131
|||||
Db 2663 CCTGTGCGGTTTCCCGACCTCTGACTTGAAGTGGCTGATTTTGTGATGCTGTCAGGGGG 2604
QY 1132 CGGAGCCTATGGAACACCCAGCAACGCGCGCGGG 1168
Db 2603 CGGAGCCTATGGAACACCCAGCAACGCGCGCGAG 2567

RESULT 13

PCT-US91-02954-12/c
; Sequence 12, Application PC/TUS9102954
; GENERAL INFORMATION:
; APPLICANT: PEPINSKY, R. BLAKE
; APPLICANT: ROSA, MARGARET D.
; APPLICANT: STOSSEL, THOMAS P.
; TITLE OF INVENTION: MULTIMERIC GELSOLIN FUSION CONSTRUCTS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02954
; FILING DATE: 19910503
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/520,368
; FILING DATE: 04-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B144CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 715-0600
; TELEFAX: (212) 715-0634
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6151 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: circular
PCT-US91-02954-12

Query Match 31.8%; Score 453.8; DB 5; Length 6151;
Best Local Similarity 99.6%; Pred. No. 1.e-132;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTACCGCACTTCAAGAACTGTAGCAGCGGCTTACATACCTCGCT 771
Db 4761 GTGTAGCCGTAGTACCGCACTTCAAGAACTGTAGCAGCGGCTTACATACCTCGCT 4702
QY 772 CTGCTAATCTGTTACAGTGGCTGCTGCAGTGGGATAAGTCTGTCTTACCGGGTTG 831

|||||
Db 4701 CTGCTAATCTGTACAGTGGCTGCTCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 4642
Qy 832 GACTCAAGACCATAGTTACCGGATAAGCGCAGCGGTGGGCTGAACCGGGGGTTCGTGC 891
Db 4641 GACTCAAGACCATAGTTACCGGATAAGCGCAGCGGTGGGCTGAACCGGGGGTTCGTGC 4582
Qy 892 ACACAGCCAGCTTTGGAGCGAACGACCTACACCGAACTGAGATACCTACACGCTGAGCAT 951
Db 4581 ACACAGCCAGCTTTGGAGCGAACGACCTACACCGAACTGAGATACCTACACGCTGAGCAT 4522
Qy 952 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 4521 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 4462
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 1071
Db 4461 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 4402
Qy 1072 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 4401 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 4342
Qy 1132 CGGAGCCTATGGAAGAAAGCGGACGAAACGCGCGCGGG 1168
Db 4341 CGGAGCCTATGGAAGAAAGCGGACGAAACGCGCGCGAG 4305

RESULT 14
US-08-460-343B-1/c
; Sequence 1, Application US/08460343B
; Patent No. 5741664
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,343B
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-mar-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-460-343B-1

Query Match 31.8%; Score 453.8; DB 1; Length 8119;

Best Local Similarity 99.6%; Pred. No. 1.5e-132;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 771
Db 2882 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 2823
Qy 772 CTGCTAATCTGTACAGTGGCTGCTCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
Db 2822 CTGCTAATCTGTACAGTGGCTGCTCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 2763
Qy 832 GACTCAAGACCATAGTTACCGGATAAGCGCAGCGGTGGGCTGAACCGGGGGTTCGTGC 891
Db 2762 GACTCAAGACCATAGTTACCGGATAAGCGCAGCGGTGGGCTGAACCGGGGGTTCGTGC 2703
Qy 892 ACACAGCCAGCTTTGGAGCGAACGACCTACACCGAACTGAGATACCTACACGCTGAGCAT 951
Db 2702 ACACAGCCAGCTTTGGAGCGAACGACCTACACCGAACTGAGATACCTACACGCTGAGCAT 2643
Qy 952 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2642 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2583
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 1071
Db 2582 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 2523
Qy 1072 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 2522 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 2463
Qy 1132 CGGAGCCTATGGAAGAAAGCGGACGAAACGCGCGCGGG 1168
Db 2462 CGGAGCCTATGGAAGAAAGCGGACGAAACGCGCGCGAG 2426

RESULT 15
US-08-398-028B-1/c
; Sequence 1, Application US/08398028B
; Patent No. 5780285
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,028B
; FILING DATE: 03-Mar-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single

TOPOLOGY: Linear
US-08-398-028B-1

Query Match		31.8%;	Score 453.8;	DB 1;	Length 8119;
Best Local Similarity		99.6%;	Pred. No. 1.5e-132;		
Matches 455;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	712	GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT	771		
Db	2882	GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT	2823		
QY	772	CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG	831		
Db	2822	CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG	2763		
QY	832	GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC	891		
Db	2762	GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC	2703		
QY	892	ACACAGCCCGAGCTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT	951		
Db	2702	ACACAGCCCGAGCTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT	2643		
QY	952	TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG	1011		
Db	2642	TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG	2583		
QY	1012	GTCGGAACAGGAGCGCACAGGGAGCTTCCAGGGGGAAACCCCTGGTATCTTTATAGT	1071		
Db	2582	GTCGGAACAGGAGCGCACAGGGAGCTTCCAGGGGGAAACCCCTGGTATCTTTATAGT	2523		
QY	1072	CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG	1131		
Db	2522	CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG	2463		
QY	1132	CGGAGCCTATGGAAAACGCCACGACGCGCGCGGG	1168		
Db	2462	CGGAGCCTATGGAAAACGCCACGACGCGCGCGAG	2426		

Search completed: January 17, 2002, 11:53:09
Job time: 17316 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: January 17, 2002, 08:54:03 ; Search time 6501.33 Seconds
(without alignments)
2355.323 Million cell updates/sec

Title: US-09-242-202a-16
Perfect score: 1425
Sequence: 1 TGCCATGGCGGGATTCTTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hic: *
10: gb_estl: *
11: gb_est2: *
12: gb_hic: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rod: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	450.2	31.6	615	10 AV735664	AV735664
C 2	449.8	31.6	527	10 AL043585	AL043585 DKF2p434G
C 3	449.8	31.6	571	10 AL044178	AL044178 DKF2p434P
C 4	449.8	31.6	579	10 AL043613	AL043613 DKF2p434H
C 5	449.8	31.6	616	10 AV735756	AV735756
C 6	449.8	31.6	617	10 AJ281661	AJ281661 4A3A-P8G1
C 7	449.8	31.6	628	11 BF381364	BF381364 ASIR0004
C 8	449.8	31.6	954	10 AL044364	AL044364 DKF2p434C
C 9	448.8	31.5	629	10 AL593919	AL593919
C 10	448.8	31.5	1004	10 AJ281480	AJ281480 4A3A-P4G8
C 11	448.2	31.5	703	10 AJ281437	AJ281437 4A3A-P4C3
C 12	448.2	31.5	800	10 AJ281449	AJ281449 4A3A-P4D5

C 13	447.8	31.4	568	10 AJ281376	AJ281376 4A3A-P2G2
C 14	447.8	31.4	579	10 AJ281320	AJ281320 4A3A-P1H1
C 15	447.8	31.4	700	10 AJ281616	AJ281616 4A3A-P8A1
C 16	446.6	31.3	1067	10 AU081137	AU081137 AU081137
C 17	445	31.2	808	10 AU176264	AU176264 AU176264
C 18	443.8	31.1	548	10 AJ281654	AJ281654 4A3A-P8F1
C 19	442.8	31.1	1070	10 AJ281552	AJ281552 4A3A-P6F1
C 20	441.8	31.0	526	10 AL043840	AL043840 DKF2p434A
C 21	441.2	31.0	498	10 AL039576	AL039576 DKF2p434D
C 22	438.8	30.8	615	10 AL044413	AL044413 DKF2p434E
C 23	438.4	30.8	741	11 BF299419	BF299419 24A-6-11
C 24	437.8	30.7	734	10 AL039459	AL039459 DKF2p434O
C 25	437.2	30.7	1089	10 AU081124	AU081124 AU081124
C 26	435.8	30.6	480	13 C111G3	C111G3
C 27	435.8	30.6	636	10 AJ281699	AJ281699 4A3A-P9E3
C 28	435.2	30.5	756	10 BE749097	BE749097 601123138
C 29	434.8	30.5	718	13 AG010489	AG010489 Homo sapi
C 30	433.6	30.4	959	10 BE749147	BE749147 601123315
C 31	431	30.2	759	10 BE749118	BE749118 601123194
C 32	429.6	30.1	840	10 BE749178	BE749178 601123444
C 33	429.2	30.1	757	11 BF863156	BF863156 963041F03
C 34	425.8	29.9	630	10 AL042334	AL042334 DKF2p434O
C 35	425.2	29.8	784	13 AQ876119	AQ876119 V133E1 mt
C 36	423.8	29.7	590	10 AV613078	AV613078 AV613078
C 37	422.6	29.7	795	13 AQ876011	AQ876011 V132B5 mt
C 38	418.6	29.4	690	13 AG009464	AG009464 Homo sapi
C 39	417.4	29.3	579	10 AV604761	AV604761 AV604761
C 40	417	29.3	540	10 AV604844	AV604844 AV604844
C 41	415.8	29.2	485	10 AL044354	AL044354 DKF2p434C
C 42	411	28.8	436	10 AL040947	AL040947 DKF2p434K
C 43	407.6	28.6	629	13 AG011367	AG011367 Homo sapi
C 44	403	28.3	794	13 AQ875988	AQ875988 V131H3 mt
C 45	401.6	28.2	706	10 BE268532	BE268532 601125429

ALIGNMENTS

RESULT 1
AV735664/c
LOCUS AV735664 615 bp mRNA EST 17-OCT-2000
DEFINITION AV735664 CB Homo sapiens CDNA clone CNAME07 5', mRNA sequence.
ACCESSION AV735664
VERSION AV735664.1 GI:10853245
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Zhang,Q., Ye,M., Wu,X., Gu,J., Huang,Q., Zhou,J., Shen,Y., Han,Z., Chen,S., Mao,M. and Chen,Z.
TITLE Homo sapiens CB library CDNA clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zhu Chen
Shanghai Institute of Hematology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64740490
Fax: 86-21-64743206
Email: mbshiem.stn.sh.cn
This clone is available at Shanghai Hematology Institute in Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
FEATURES
source
1. 615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CNAME07"
/clone_lib="CB"
/tissue_type="cord blood"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"

/note="Vector: pBluescript; Site_1: EcoRI; The insert is
Cloned randomly with the EcoRI digestion"
BASE COUNT 129 a 176 c 159 g 147 t 4 others
ORIGIN

Query Match 31.6%; Score 450.2; DB 10; Length 615;
Best Local Similarity 99.1%; Pred. No. 2.7e-116;
Matches 452; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCCGTACATACCTCGCT 771
Db 460 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCCGTACATACCTCGCT 771
QY 772 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGGATAGTCTGTACCGGGTTG 831
Db 400 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGGATAGTCTGTACCGGGTTG 341
QY 832 GACTCAAGACGATAGTTACCGGATAGGCGCAGCGTTCGGGCTGAACGGGGTTCGTGC 891
Db 340 GACTCAAGACGATAGTTACCGGATAGGCGCAGCGTTCGGGCTGAACGGGGTTCGTGC 281
QY 892 ACACAGCCAGCTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCGTGCAGCAT 951
Db 280 ACACAGCCAGCTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCGTGCAGCAT 221
QY 952 TGAGAAAGCCCACTGCTTCCGGAAGGGAAGAGCGGAGTATCCGGTAAAGCGGCGAGG 1011
Db 220 TGAGAAAGCCCACTGCTTCCGGAAGGGAAGAGCGGAGTATCCGGTAAAGCGGCGAGG 161
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCTGTGTATCTTTATAGT 1071
Db 160 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCTGTGTATCTTTATAGT 101
QY 1072 CTTGTCCGGTTTCGCCACCTCTGACTTCAGCGTCGATTTTGTGATCTCTCAGGGGGG 1131
Db 100 CTTGTCCGGTTTCGCCACCTCTGACTTCAGCGTCGATTTTGTGATCTCTCAGGGGGG 41
QY 1132 CGAGCGCTATGAAAGAACCGCAACGCGGCGCGG 1167
Db 40 CGAGCGCTATGAAAGAACCGCAACGCGGCGCGG 5

RESULT 2
AL043585/c
LOCUS
DEFINITION DKFZP434G0127_s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL043585
VERSION AL043585
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 527)
AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 8, 1999 this sequence version replaced gi:5866785.
Contact: Bloecker H
MIPS

An Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZP434G0127) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

FEATURES
Source Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..527

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZP434G0127"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 113 a 158 c 137 g 119 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 527;
Best Local Similarity 99.6%; Pred. No. 3.3e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCCGTACATACCTCGCT 771
Db 474 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCCGTACATACCTCGCT 771
QY 772 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGGATAGTCTGTACCGGGTTG 831
Db 414 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGGATAGTCTGTACCGGGTTG 355
QY 832 GACTCAAGACGATAGTTACCGGATAGGCGCAGCGTTCGGGCTGAACGGGGTTCGTGC 891
Db 354 GACTCAAGACGATAGTTACCGGATAGGCGCAGCGTTCGGGCTGAACGGGGTTCGTGC 295
QY 892 ACACAGCCCACTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCGTGCAGCAT 951
Db 294 ACACAGCCCACTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCGTGCAGCAT 235
QY 952 TGAGAAAGCCCACTGCTTCCGGAAGGGAAGAGCGGAGTATCCGGTAAAGCGGCGAGG 1011
Db 234 TGAGAAAGCCCACTGCTTCCGGAAGGGAAGAGCGGAGTATCCGGTAAAGCGGCGAGG 175
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCTGTGTATCTTTATAGT 1071
Db 174 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCTGTGTATCTTTATAGT 115
QY 1072 CTTGTCCGGTTTCGCCACCTCTGACTTCAGCGTCGATTTTGTGATCTCTCAGGGGGG 1131
Db 114 CTTGTCCGGTTTCGCCACCTCTGACTTCAGCGTCGATTTTGTGATCTCTCAGGGGGG 55
QY 1132 CGAGCGCTATGAAAGAACCGCAACGCGGCGC 1164
Db 54 CGAGCGCTATGAAAGAACCGCAACGCGGCGC 22

RESULT 3
AL044178/c
LOCUS
DEFINITION DKFZP434P0828_s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL044178
VERSION AL044178
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 571)
AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 9, 1999 this sequence version replaced gi:5866789.
Contact: Bloecker H
MIPS

An Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZP434P0828) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.

No r1 sequence available.
This clone (DKFZp434P0828) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1. .571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434P0828"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 122 a 168 c 149 g 132 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 571;
Best Local Similarity 99.6%; Pred. No. 3.4e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCCACACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 771
DB 463 GTGTAGCCGTAGTTAGGCCACACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 404
QY 772 CTGCTAATCTCTGTACCTGCTGCTGCAGTGGCGGATAAGTCGTCCTTACCGGGTTG 831
DB 403 CTGCTAATCTCTGTACCTGCTGCTGCAGTGGCGGATAAGTCGTCCTTACCGGGTTG 344
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC 891
DB 343 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC 284
QY 892 ACACAGCCAGCTTGGAGCGACGACCTACACGGAAGTATACCTACAGCTGAGCAT 951
DB 283 ACACAGCCAGCTTGGAGCGACGACCTACACGGAAGTATACCTACAGCTGAGCAT 224
QY 952 TGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 223 TGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 164
QY 1012 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGATCTTTATAGT 1071
DB 163 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGATCTTTATAGT 104
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTCAGGGGG 1131
DB 103 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTCAGGGGG 44
QY 1132 CGGAGCCTATGGAAGAAACGCCAGCAACGCGGCC 1164
DB 43 CGGAGCCTATGGAAGAAACGCCAGCAACGCGGCC 11

RESULT 4
AL043613/c
LOCUS
DEFINITION DKFZp434H1527_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION. AL043613
VERSION AL043613.1 GI:5423000
KEYWORDS EST..
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 579)
Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Bloecker, et al.)
Unpublished (1999)
Contact: Bloecker H
MIPS

An Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.

No r1 sequence available.
This clone (DKFZp434H1527) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1. .579
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434H1527"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 128 a 171 c 152 g 128 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 579;
Best Local Similarity 99.6%; Pred. No. 3.4e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCCACACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 771
DB 477 GTGTAGCCGTAGTTAGGCCACACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 418
QY 772 CTGCTAATCTCTGTACCTGCTGCTGCAGTGGCGGATAAGTCGTCCTTACCGGGTTG 831
DB 417 CTGCTAATCTCTGTACCTGCTGCTGCAGTGGCGGATAAGTCGTCCTTACCGGGTTG 358
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC 891
DB 357 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC 298
QY 892 ACACAGCCAGCTTGGAGCGACGACCTACACGGAAGTATACCTACAGCTGAGCAT 951
DB 297 ACACAGCCAGCTTGGAGCGACGACCTACACGGAAGTATACCTACAGCTGAGCAT 238
QY 952 TGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 237 TGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 178
QY 1012 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGATCTTTATAGT 1071
DB 177 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGATCTTTATAGT 118
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTCAGGGGG 1131
DB 117 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTCAGGGGG 58
QY 1132 CGGAGCCTATGGAAGAAACGCCAGCAACGCGGCC 1164
DB 57 CGGAGCCTATGGAAGAAACGCCAGCAACGCGGCC 25

RESULT 5
AV735756/c
LOCUS
DEFINITION AV735756 CB Homo sapiens cDNA clone CBMAGC03 5', mRNA sequence.

RESULT	6
AJ281661/c	
LOCUS	617 bp mRNA EST 30-JUN-2000
DEFINITION	4A3A-P8G10-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P8G10, mRNA sequence.
ACCESSION	AJ281661
VERSION	AJ281661.1 GI:6929540
KEYWORDS	EST;
SOURCE	African malaria mosquito.
ORGANISM	Anopheles gambiae
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
AUTHORS	Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C., Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B. and Kafatos,F.C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE	20300950
COMMENT	Contact: Dimopoulos G Potis C. Kafatos laboratory European Molecular Biology Laboratory Meyerohofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1. 617 /organism="Anopheles gambiae" /strain="4A r/r" /db_xref="taxon:7165" /clone="4A3A-P8G10" /cdone="4A3A-P8G10" /cell_line="Immune competent 4A3A" /lab_host="E. coli DH10B" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996); Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT	136 A 175 C 162 G 144 T
ORIGIN	
Query Match	31.6%; Score 449.8; DB 10; Length 617;
Best Local Similarity	99.6%; Pred. No. 3.5e-116;
Matches	451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	712 GTGTAGCCGTAGTTAGGCCACCACCTTCAAGAAGCTCTGTAGCACCGCCTACATACCTCGCT 771
Db	
Qy	488 GTGTAGCCGTAGTTAGGCCACCACCTTCAAGAAGCTCTGTAGCACCGCCTACATACCTCGCT 429
Db	
Qy	772 CTGCTAATCTCTTACCAGTGCTCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 831
Db	
Qy	428 CTGCTAATCTCTTACCAGTGCTCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 369
Db	
Qy	832 GACTCAAGACGATATTACCGGTAAGCGGCAGCGGTCGGGCTCAACGGGGGTTCTGTCG 891
Db	
Qy	368 GACTCAAGACGATATTACCGGTAAGCGGCAGCGGTCGGGCTCAACGGGGGTTCTGTCG 309
Db	
Qy	892 ACACAGCCAGCTTGGAGCGAACACCTCACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db	
Qy	308 ACACAGCCAGCTTGGAGCGAACACCTCACCGAACTGAGATACCTACAGCGTGAGCTA 249
Db	
Qy	952 TCAGAAAGCGCCAGCTTCCCAGTAGGAGGAAGCGGCAGAGTATCCCGGTAGCGGCAGG 1011
Db	
Qy	248 TCAGAAAGCGCCAGCTTCCCAGTAGGAGGAAGCGGCAGAGTATCCCGGTAGCGGCAGG 189
Db	
Qy	1012 CTCGGAACGAGAGCGCACAGGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 1071
Db	


```
|||||
Db 188 GTCGAACAGGAGCGCAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 129
|||||
QY 1072 CCGTCCGGGTTCCGACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 128 CCGTCCGGGTTCCGACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 69
|||||
QY 1132 CGGAGCCTATGGAAGAACGCGCAACGCGGCC 1164
|||||
Db 68 CGGAGCCTATGGAAGAACGCGCAACGCGGCC 36
|||||

RESULT 7
BF381364/c 628 bp mRNA EST 27-NOV-2000
LOCUS
DEFINITION AsIR0004 Mosquito Genes Pool Related Malaria Infection Anopheles
stephensi cDNA 5', mRNA sequence.
ACCESSION BF381364
VERSION BF381364.1 GI:11370487
KEYWORDS EST.
SOURCE Anopheles stephensi.
ORGANISM Anopheles stephensi.
REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoides
; Anopheles.
TITLE 1 (bases 1 to 628)
XU, X., QU, F., SONG, G. and XU, J.
The differentially expressing genes pool from Anopheles stephensi
related to infection with Plasmodium yoelii enriched by suppression
subtractive hybridization
JOURNAL Unpublished (2001)
COMMENT Contact: Xu Xiaochun; Qu Fengyi; Song Guanhong; Xu Jiannong
Department of Parasitology
Second Military Medical University
800 Xiangyin Rd., Shanghai, 200433, China
Tel: 86 021 25070276
Email: xcxu@smmu.edu.cn
Seq primer: T7
High quality sequence stop: 629
POLYA=No. Location/Qualifiers
source 1..628
/organism="Anopheles stephensi"
/strain="Hof"
/db_xref="taxon:30069"
/clone_lib="Mosquito Genes Pool Related Malaria Infection"
/sex="female"
/tissue_type="whole body"
/dev_stage="24 hours post-infection"
BASE COUNT 145 a 178 c 169 g 136 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 11; Length 628;
Best Local Similarity 99.6%; Pred. No. 3.5e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
|||||
Db 569 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 510
|||||
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCGCCAGTGGCGGATAAGTCTCTTACCGGGTTG 831
|||||
Db 509 CTGCTAATCTGTTACCAAGTGGCTGCGCCAGTGGCGGATAAGTCTCTTACCGGGTTG 450
|||||
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGGCTGAACCGGGGTTCTGTC 891
|||||
Db 449 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGGCTGAACCGGGGTTCTGTC 390
|||||
QY 892 ACACAGCCCGAGTTGGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGGTGACAT 951
|||||
Db 389 ACACAGCCCGAGTTGGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGGTGACAT 932
|||||
```

```
QY 952 TGAAAGACCCACCGCTTCCGAAGGAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCGAG 1011
|||||
Db 329 TGAAAGACCCACCGCTTCCGAAGGAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCGAG 270
|||||
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 1071
|||||
Db 269 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTCGTATCTTTATAGT 210
|||||
QY 1072 CCGTCCGGGTTCCGACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 209 CCGTCCGGGTTCCGACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 150
|||||
QY 1132 CGGAGCCTATGGAAGAACGCGCAACGCGGCC 1164
|||||
Db 149 CGGAGCCTATGGAAGAACGCGCAACGCGGCC 117
|||||

RESULT 8
AL044364/c 954 bp mRNA EST 29-FEB-2000
LOCUS
DEFINITION DKFZp434C172.sl 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL044364
VERSION AL044364.1 GI:5432586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 954)
XANSORGE, W., BENES, V., KRIEGER, S., MEWES, H. W., GASSENHUBER, J. and
WIEMANN, S.
EST (Ansoerge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansoerge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source 1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="DKFZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 954;
Best Local Similarity 99.6%; Pred. No. 4e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
|||||
Db 851 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 792
|||||
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
|||||
Db 791 CTGCTAATCTGTTACCAAGTGGCTGCGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 732
|||||
```

```

OY 832 GACTCAAGCAGTATGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGTTCTGTC 891
Db 731 GACTCAAGCAGTATGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGTTCTGTC 672
OY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 671 ACACAGCCAGCTTGGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCGTGAGCAT 612
OY 952 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGG 1011
Db 611 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGG 552
OY 1012 GTCGGAACAGGAGCGCAGCGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGT 1071
Db 551 GTCGGAACAGGAGCGCAGCGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGT 492
OY 1072 CTTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTCTGATGCTCGTCAGGGGGG 1131
Db 491 CTTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTCTGATGCTCGTCAGGGGGG 432
OY 1132 CGGAGCTATGAAAGACCGCAGCAGCGGCC 1164
Db 431 CGGAGCTATGAAAGACCGCAGCAGCGGCC 399

RESULT 9
AL593919/c 629 bp mRNA EST 30-JUL-2001
LOCUS AL593919 XGC-gastrula silurana tropicalis cDNA clone Tgas003010 5',
DEFINITION mRNA sequence.
ACCESSION AL593919
VERSION AL593919.1 GI:15005980
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 629)
Huckle, E., Taylor, R., McMurray, A., Ashurst, J. L., Zorn, A. M. and
Rogers, J.
Sanger Xenopus tropicalis EST project 2001
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinnton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tgas003010.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
FEATURES
    source
        1..629
        /organism="Silurana tropicalis"
        /db_xref="taxon:8364"
        /clone="Tgas003010"
        /clone_lib="XGC-gastrula"
        /dev_stage="gastrula (stages 10.5-13 mixed)"
        /lab_host="Escherichia coli DH10B"
        /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
        was oligo dt primed from 5ug of poly A+ RNA from stages
        10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
        into pCS107 with EcoRI at the 5' end and NotI at the 3'
        end."
BASE COUNT 149 a 181 c 164 g 135 t
ORIGIN

Query Match 31.5%; Score 448.8; DB 10; Length 629;
Best Local Similarity 99.6%; Pred. No. 6.7e-116;
Matches 450; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 712 GTGTAGCCGTAGTTAGCGCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
Db 528 GTGTAGCCGTAGTTAGCGCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 469
OY 772 CTGCTAATCTCTGTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTG 831
Db 468 CTGCTAATCTCTGTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTG 409
OY 832 GACTCAAGCAGTATGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGTTCTGTC 891
Db 408 GACTCAAGCAGTATGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGTTCTGTC 349
OY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 348 ACACAGCCAGCTTGGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCGTGAGCAT 289
OY 952 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGG 1011
Db 288 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGG 229
OY 1012 GTCGGAACAGGAGCGCAGCGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGT 1071
Db 228 GTCGGAACAGGAGCGCAGCGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGT 169
OY 1072 CTTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTCTGATGCTCGTCAGGGGGG 1131
Db 168 CTTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTCTGATGCTCGTCAGGGGGG 109
OY 1132 CGGAGCTATGAAAGACCGCAGCAGCGGCC 1163
Db 108 CGGAGCTATGAAAGACCGCAGCAGCGGCC 77

RESULT 10
AL593919/c 1004 bp mRNA EST 30-JUN-2000
LOCUS AL593919 XGC-gastrula silurana tropicalis cDNA clone Tgas003010 5',
DEFINITION mRNA sequence.
ACCESSION AL593919
VERSION AL593919.1 GI:6929360
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
1 (bases 1 to 1004)
Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M. B.
and Kafatos, F. C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
    1..1004
    /organism="Anopheles gambiae"
    /strain="4A r/r"
    /db_xref="taxon:7165"
    /clone="4A3A-P4G8"
    /clone_lib="Anopheles gambiae immune competent 4A3A"
    /lab_host="E. coli DH10B"
    /note="Vector: pT73D-Pac (Pharmacia) with a modified
    polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
    forward priming site which reads from the 3' end of the
    cDNA. The 4A3A is a directionally cloned and normalized
    cDNA library that was constructed from the 4A3A cell line

```

oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others

Query Match 31.5%; Score 448.8; DB 10; Length 1004;
Best Local Similarity 99.3%; Pred. No. 7.8e-116;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
DB 475 GTGTAGCGGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 416
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 831
DB 415 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 356
QY 832 GACTCAAGCAGATGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 891
DB 355 GACTCAAGCAGATGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 296
QY 892 ACACAGCCCGAGTTGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 295 ANACAGCCCGAGTTGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGTGAGCTA 236
QY 952 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 235 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 176
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 1071
DB 175 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 116
QY 1072 CTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
DB 115 CTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 56
QY 1132 CGGAGCCTATGGAAGAAACGCGCAACGCGGCC 1164
DB 55 CGGAGCCTATGGAAGAAACGCGCAACGCGGCC 23

RESULT 11
AJ281437/c
LOCUS 703 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-P4C3-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4C3, mRNA sequence.

ACCESSION AJ281437
VERSION AJ281437.1 GI:6929317
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae

REFERENCE 1 (bases 1 to 703)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C., Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoorge,W., Soares,M.B. and Kafatos,F.C.

TITLE mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .703
/organism="Anopheles gambiae"

/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4C3"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 159 a 204 c 193 g 147 t
ORIGIN

Query Match 31.5%; Score 448.2; DB 10; Length 703;
Best Local Similarity 99.3%; Pred. No. 1e-115;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
DB 692 GTGTAGCGGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 633
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 831
DB 632 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 573
QY 832 GACTCAAGCAGATGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 891
DB 572 GACTCAAGCAGATGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 513
QY 892 ACACAGCCCGAGTTGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 512 ACACAGCCCGAGTTGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGTGAGCTA 453
QY 952 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 452 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 393
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 1071
DB 392 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 333
QY 1072 CTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
DB 332 CTGTGCGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 273
QY 1132 CGGAGCCTATGGAAGAAACGCGCAACGCGGCC 1164
DB 272 CGGAGCCTATGGAAGAAACGCGCAACGCGGCC 240

RESULT 12
AJ281449/c
LOCUS 800 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4D5, mRNA sequence.

ACCESSION AJ281449
VERSION AJ281449.1 GI:6929329
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae

REFERENCE 1 (bases 1 to 800)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C., Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoorge,W., Soares,M.B. and Kafatos,F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of

mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. 800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P405"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 203 a 206 c 198 g 193 t
ORIGIN

Query Match 31.5%; Score 448.2; DB 10; Length 800;
Best Local Similarity 99.3%; Pred. No. 1.1e-115;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTACCGTAGTGGCCACCACCTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 771
|||||
Db 502 GTGTACCGTAGTGGCCACCACCTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 771
|||||
QY 772 CTGCTAATCTGTACAGTGGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
|||||
Db 442 CTGCTAATCTGTACAGTGGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
|||||
QY 832 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGTTCTGTCG 891
|||||
Db 382 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGTTCTGTCG 891
|||||
QY 892 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTACAGTACCTACAGCTGAGCTA 263
|||||
Db 322 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTACAGTACCTACAGCTGAGCTA 263
|||||
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGGAGGAGTATCCGGTAAGCGCGAGG 1011
|||||
Db 262 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGGAGGAGTATCCGGTAAGCGCGAGG 203
|||||
QY 1012 GTCGGAACGAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCGGAGTATCCGGTAAGCGCGAGG 1071
|||||
Db 202 GTCGGAACGAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCGGAGTATCCGGTAAGCGCGAGG 1131
|||||
QY 1072 CCGTCCGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAAGGGGG 143
|||||
Db 142 CCGTCCGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAAGGGGG 83
|||||
QY 1132 CGGAGCCTATGAAAAAGCCAGCAACCGCGCC 1164
|||||
Db 82 CGGAGCCTATGAAAAAGCCAGCAACCGCGCC 50
|||||

RESULT 13
AJ281376/c
LOCUS
DEFINITION 4A3A-P2G2-F Anopheles gambiae Immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P2G2, mRNA sequence.
ACCESSION AJ281376
VERSION AJ281376.1 GI:6929257

EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles
1 (bases 1 to 568)
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. 568
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P2G2"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 123 a 167 c 149 g 127 t
ORIGIN

Query Match 31.4%; Score 447.8; DB 10; Length 568;
Best Local Similarity 99.1%; Pred. No. 1.2e-115;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTGGCCACCACCTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 771
|||||
Db 466 GTGTAGCCGTAGTGGCCACCACCTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 407
|||||
QY 772 CTGCTAATCTGTACAGTGGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
|||||
Db 406 CTGCTAATCTGTACAGTGGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 347
|||||
QY 832 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGTTCTGTCG 891
|||||
Db 346 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGTTCTGTCG 287
|||||
QY 892 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTACAGTACCTACAGCTGAGCAT 951
|||||
Db 286 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTACAGTACCTACAGCTGAGCAT 227
|||||
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGGAGGAGTATCCGGTAAGCGCGAGG 1011
|||||
Db 226 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGGAGGAGTATCCGGTAAGCGCGAGG 167
|||||
QY 1012 GTCGGAACGAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCGGAGTATCTTTATAGT 1071
|||||
Db 166 GTCGGAACGAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCGGAGTATCTTTATAGT 107
|||||
QY 1072 CCTGTCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 106 CCTGTCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 47
|||||
QY 1132 CGGAGCCTATGAAAAAGCCAGCAACCGCGCC 1164
|||||

```

|||||
Db 46 CGGAGCCTATGGAAGCGCCNCCACGCGCC 14

RESULT 14
AJ281320/c 579 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-PIH1-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-PIH1, mRNA sequence.
ACCESSION AJ281320
VERSION AJ281320.1 GI:6929201
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
REFERENCE 1 (bases 1 to 579)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerothofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .579
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone_lib="4A3A-PIH1"
/cell_line="Anopheles gambiae immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 120 a 167 c 151 g 137 t 4 others
ORIGIN

Query Match 31.4%; Score 447.8; DB 10; Length 579;
Best Local Similarity 99.1%; Pred. No. 1.2e-115;
Matches 449; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTCTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
|||||
Db 456 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 397
|||||
QY 772 CTGCTAATCCTGTTACCAAGTGGCTGCTGCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 831
|||||
Db 396 CTGCTAATCCTGTTACCAAGTGGCTGCTGCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 337
|||||
QY 832 GACTCAAGCAGTATGTTACCGGATAGAGCGGCGAGTGGCTGCGGCTGAACGGGGGTTGTCG 891
|||||
Db 336 GACTCAAGCAGTATGTTACCGGATAGAGCGGCGAGTGGCTGCGGCTGAACGGGGGTTGTCG 277
|||||
QY 892 ACACGCCAGCTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCGTGAACAT 951
|||||
Db 276 ACACGCCAGCTTGGAGCGAAGCACTACACGAACTGAGATACCTACAGCGTGAACAT 217
|||||
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAAAGCGGAGG 1011
|||||

```

```

Db 216 TGAGAAAGCGCCACGCTTCCGAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGAGG 157
QY 1012 GTCGGAACAGGAGAGCGGACGAGGAGCTTCCAGGGGGAACGCTCGTATCTTTATAGT 1071
|||||
Db 156 GTCGGAACAGGAGAGCGGACGAGGAGCTTCCAGGGGGAACGCTCGTATCTTTATAGT 97
QY 1072 CTTGTCTGGGTTTCCGCCACCTCTGACTTGTAGGTCGCGATTTTGTGATCTCTCGTCAGGGGG 1131
|||||
Db 96 CTTGTCTGGGTTTCCGCCACCTCTGASGTGAGCGTGTGATTTTGTGATCTCTCGTCAGGGGG 37
QY 1132 CGGAGCCTATGGAAGCGCCACGAGCAACGCGGCC 1164
|||||
Db 36 CGGAGCCTATGGAAGCGCCACGAGCAACGCGGCC 4

RESULT 15
AJ281616/c 700 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P8A12-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P8A12, mRNA sequence.
ACCESSION AJ281616
VERSION AJ281616.1 GI:6929495
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
REFERENCE 1 (bases 1 to 700)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerothofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .700
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone_lib="4A3A-P8A12"
/cell_line="Anopheles gambiae immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 167 a 196 c 189 g 146 t 2 others
ORIGIN

Query Match 31.4%; Score 447.8; DB 10; Length 700;
Best Local Similarity 99.1%; Pred. No. 1.3e-115;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
|||||
Db 599 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 540
|||||
QY 772 CTGCTAATCCTGTTACCAAGTGGCTGCTGCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 831
|||||
Db 539 CTGCTAATCCTGTTACCAAGTGGCTGCTGCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 480

```

Search completed: January 17, 2002, 08:54:11
Job time: 7123 sec

source		1. .3768		/organism="Homo sapiens"		/db_xref="taxon:9606"		719 t	
BASE COUNT		758 a		1170 c		1121 g		719 t	
ORIGIN									
Query Match		31.3%;		Score 666;		DB 6;		Length 3768;	
Best Local Similarity		98.4%;		Pred. No. 1.2e-154;					
Matches		683;		Conservative		0;		Mismatches 10; Indels 1; Gaps 1;	
QY	228	ACGGCAGCAGAA	GATCACA	TGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACCA	287				
Db	3075	ACCCACAGCAGGG	CTTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACCA	3134					
QY	288	CAGGCACCGCAGC	ATCATCTATCCAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCC	347					
Db	3135	CAGGCACCGCAGC	ATCATCTATCCAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCC	3194					
QY	348	CTCTGAAGAGGAGG	CCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGT	407					
Db	3195	CTCTGAAGAGGAGG	CCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGT	3254					
QY	408	ATTTGATGGTGAC	CTGGGAATGGGGCAGCCCAAGGGGCTGCAAGCCCTCCCCACACATGA	467					
Db	3255	ATTTGATGGTGAC	CTGGGAATGGGGCAGCCCAAGGGGCTGCAAGCCCTCCCCACACATGA	3314					
QY	468	CCCGAGCCCTCTAC	AGCGGTACAGTGGAGCCCAAGTACCCCTGCTGAGACTGA	527					
Db	3315	CCCGAGCCCTCTAC	AGCGGTACAGTGGAGCCCAAGTACCCCTGCTGAGACTGA	3374					
QY	528	TGGCTACGTTGCC	CCCCCTGACCTGCAGCCCCCAGCCTGAAATATGTGAACCCAGCAGATGT	587					
Db	3375	TGGCTACGTTGCC	CCCCCTGACCTGCAGCCCCCAGCCTGAAATATGTGAACCCAGCAGATGT	3434					
QY	588	TCGGCCCCAGCCCT	TCGCCCGAGAGGGCCCTCTGCCTGCTGCCCGACCTGCTGTGTC	647					
Db	3435	TCGGCCCCAGCCCT	TCGCCCGAGAGGGCCCTCTGCCTGCTGCCCGACCTGCTGTGTC	3494					
QY	648	CACCTCTGGAAGG	CCCCAAGACTCTCTCCCAGGGAAGATGGGTCGTCAAGACGTTTT	707					
Db	3495	CACCTCTGGAAGG	CCCCAAGACTCTCTCCCAGGGAAGATGGGTCGTCAAGACGTTTT	3554					
QY	708	TGCTTTGGGGTG	CGCTGAGAACCCCGAG--ACTTGACACCCCGAGGAGGAGTGCCCC	766					
Db	3555	TGCTTTGGGGTG	CGCTGAGAACCCCGAG--ACTTGACACCCCGAGGAGGAGTGCCCC	3614					
QY	767	TCAGCCCCACCTCT	CTCTGCTTCAGCCCGACCTTCGACAACTCTATTACTGGGACCA	826					
Db	3615	TCAGCCCCACCTCT	CTCTGCTTCAGCCCGACCTTCGACAACTCTATTACTGGGACCA	3674					
QY	827	GGACCCACAGAG	CGGGGCTCCACCCAGCACCCTCAAGGGACACCTACGGCAGAGAA	886					
Db	3675	GGACCCACAGAG	CGGGGCTCCACCCAGCACCCTCAAGGGGACACCTACGGCAGAGAA	3734					
QY	887	CCCAGAGTACCT	GGGTCTGGACGTGCCAGTGTGA	920					
Db	3735	CCCAGAGTACCT	GGGTCTGGACGTGCCAGTGTGA	3768					
RESULT 9									
AB008451		3780 bp mRNA							
LOCUS		Canis familiaris mRNA for erbB-2, complete cds.							
DEFINITION		AB008451							
ACCESSION		AB008451.1 GI:2575866							
VERSION		erbB-2							
KEYWORDS		Canis familiaris cdna to mRNA.							
SOURCE		Canis familiaris							
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.							
REFERENCE		1 (bases 1 to 3780)							
AUTHORS		Yokota, H.							

Query Match		31.4%;		Score 667.6;		DB 6;		Length 3768;	
Best Local Similarity		98.6%;		Pred. No. 4.8e-155;					
Matches		684;		Conservative		0;		Mismatches 9; Indels 1; Gaps 1;	
QY	228	ACGGCAGCAGAA	GATCACA	TGTCACAGACCTGCCCGGGCGCTGGGGGCATGGTCCACCA	287				
Db	3075	ACCCACAGCAGGG	CTTCTTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACCA	3134					
QY	288	CAGGCACCGCAGC	ATCATCTATCCAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCC	347					
Db	3135	CAGGCACCGCAGC	ATCATCTATCCAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCC	3194					
QY	348	CTCTGAAGAGGAGG	CCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGT	407					
Db	3195	CTCTGAAGAGGAGG	CCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGT	3254					
QY	408	ATTTGATGGTGAC	CTGGGAATGGGGCAGCCCAAGGGGCTGCAAGCCCTCCCCACACATGA	467					
Db	3255	ATTTGATGGTGAC	CTGGGAATGGGGCAGCCCAAGGGGCTGCAAGCCCTCCCCACACATGA	3314					
QY	468	CCCGAGCCCTCTAC	AGCGGTACAGTGGAGCCCAAGTACCCCTGCTGAGACTGA	527					
Db	3315	CCCGAGCCCTCTAC	AGCGGTACAGTGGAGCCCAAGTACCCCTGCTGAGACTGA	3374					
QY	528	TGGCTACGTTGCC	CCCCCTGACCTGCAGCCCCCAGCCTGAAATATGTGAACCCAGCAGATGT	587					
Db	3375	TGGCTACGTTGCC	CCCCCTGACCTGCAGCCCCCAGCCTGAAATATGTGAACCCAGCAGATGT	3434					
QY	588	TCGGCCCCAGCCCT	TCGCCCGAGAGGGCCCTCTGCCTGCTGCCCGACCTGCTGTGTC	647					
Db	3435	TCGGCCCCAGCCCT	TCGCCCGAGAGGGCCCTCTGCCTGCTGCCCGACCTGCTGTGTC	3494					
QY	648	CACCTCTGGAAGG	CCCCAAGACTCTCTCCCAGGGAAGATGGGTCGTCAAGACGTTTT	707					
Db	3495	CACCTCTGGAAGG	CCCCAAGACTCTCTCCCAGGGAAGATGGGTCGTCAAGACGTTTT	3554					
QY	708	TGCTTTGGGGTG	CGCTGAGAACCCCGAG--ACTTGACACCCCGAGGAGGAGTGCCCC	766					
Db	3555	TGCTTTGGGGTG	CGCTGAGAACCCCGAG--ACTTGACACCCCGAGGAGGAGTGCCCC	3614					
QY	767	TCAGCCCCACCTCT	CTCTGCTTCAGCCCGACCTTCGACAACTCTATTACTGGGACCA	826					
Db	3615	TCAGCCCCACCTCT	CTCTGCTTCAGCCCGACCTTCGACAACTCTATTACTGGGACCA	3674					
QY	827	GGACCCACAGAG	CGGGGCTCCACCCAGCACCCTCAAGGGACACCTACGGCAGAGAA	886					
Db	3675	GGACCCACAGAG	CGGGGCTCCACCCAGCACCCTCAAGGGGACACCTACGGCAGAGAA	3734					
QY	887	CCCAGAGTACCT	GGGTCTGGACGTGCCAGTGTGA	920					
Db	3735	CCCAGAGTACCT	GGGTCTGGACGTGCCAGTGTGA	3768					
RESULT 8									
AX060704		3768 bp DNA							
LOCUS		Sequence 2 from Patent WO0100244.							
DEFINITION		AX060704							
ACCESSION		AX060704.1 GI:12406101							
VERSION		human.							
KEYWORDS		Homo sapiens							
SOURCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
ORGANISM		1 (bases 1 to 3768)							
REFERENCE		Erickson, S. and Schwall, R.							
AUTHORS		Methods of treatment using anti-erbB antibody-maytansinoid							
TITLE		conjugates							
JOURNAL		Patent: WO 0100244-A 2 04-JAN-2001;							
FEATURES		Genentech, Inc. (US)							
		Location/Qualifiers							

Direct Submission
Submitted (23-OCT-1997) to the DDBJ/EMBL/GenBank databases. Hiroshi Yokota, Rakuno Gakuen University, Veterinary Biochemistry, Bunkyo-daimeidomari, Ebetsu, Hokkaido 069, Japan (E-mail: BXA03503@niftyserve.or.jp, Tel: 011-386-1111, Fax: 011-387-5890)
2 (bases 1 to 3780)
Yokota, H.
cDNA cloning of erbB-2 from canine mammary gland
Published Only in Database (1997) In press
Location/Qualifiers
1. 3780
/organism="Canis familiaris"
/db_xref="taxon:9615"
1. 3780
/codon_start=1
/product="erbB-2"
/protein_id="BAA33127.1"
/db_xref="GI:2575867"
/translation="MELAAWCRWGLLLALLPSGAGTQVCTGDMKRLRLPASPTHLD MLRLHYGCOVVOGNLELTLYPANASLFLQIOEVGYVLIASHOVROIPLORLRV RGTOLFEDNVALVLDNGDPLEGGIPAPGAQGLRELQLSLTELKGVLIQSRQ LCHQDTLWKDVFHKNQALTLIDTRNFSACPPCSPPACKDAHGWAGSSGDCQSLRT VAGGARCPQPTDCCHQCAAGCTGPKHSDCLACHFNHSGICELHCPALTYNT DTFESMPNPERYTFGASCVTSCPNYLDSTGCTLVCLPNNQEVTAEDGTQRCRK SKPCARVYGLGMEHLREAVTSANIQEFAGCKKIFGSLAFLPSDGDGPASNTAPL OPEOLRYFEALIEETLYLISAWPDSLNLVSFONLRVIRGRVLHDGAYSLTLOGLI SWGLRSIRLGLSGALILHNRALCFVHTVPMDQLFRNQHALLHSANRPERECVGG LACYPCHAGHCWGPPTQCVNSQFLRQECVEECRVLQGLPREYIKWDFCLPCHSEC QPQNGVTCGSEADQVACAHYKDPFCVARCPGSKVPDLSEMPKWFADDEETCP CPINRTHSCADLDKGPAPQASPVTSIAAVVGIILAVVGLVGLILIKRRQKIR KYTMRLLQETELVETLPSGAMPNQMRILKTELKVKVGLGAGFGVYKGIWIP DMENVKIPVAIKVLRKNTSPKANKEILDEAVMAGVSPVYSLILGLICTSTVOLVTO LMPVGLLDHVRHGRSLQDILLWCQVIAKMSYLEDVRLVHDLAARNVLYKSPN HVKITDGLARLLDIDETHADGKVPKIMMALESIPPRFTHQSDVWSVGTWVWEL MTFGAKYDGPAREIPDLLEKGERLPOPICTIDVIMVMKWMIDSECRPRELY AEFERMDARQFVIEIDLGPAISPLDSTFYRSLDEDDMDGLVDAEYGLVPOQGF CPEPTGAGGTAAHRHRSSTNRNGGELTGLEPSEEPKSPPLASPEGAGSDVDFGD LGMAAAGLSDPSQPLQRYSEDVTPPLPPTDKVAPLTCSPQEVVNOPEVWP OPPLALRGLPLPPSPACATLERPKTLSPKNGVVKVDVAFGSAVENPEYLAPR GRAAPQHPHPAPSPAFDNLVYWDQDSERGSPPSTFEGIPTAENPEYLGUDVPV"

BASE COUNT 743 a 1162 c 1150 g 725 t
ORIGIN

Query Match 24.3%; Score 516.2; DB 4; Length 3780;
Best Local Similarity 85.3%; Pred. No. 1.8e-117;
Matches 605; Conservative 0; Mismatches 88; Indels 16; Gaps 2;

Qy 228 ACGGCAGCAGAGATGCATGTCACAGACCCCTGCCCGGCGCTGGGGCATGGTCCACCA 287
Db 3072 ACCCCAGCAGGGTTCTTCTGCCACAGACCTACCCAGGGCGTGGGGCATGCCACCG 3131

Qy 288 CAGGCACCGCAGCTCATCTACAGAGGTGGGGTGGGACCTTGACATGAGGCTGGAGCC 347
Db 3132 ACGGCACCGCAGCTCATCAGCAGGAATGGGGTGGTGGTGGTGGTGGTGGTGGTGG 3191

Qy 348 CTCTCAAGAGAGGAGGCGCCAGCTCCACTGGCACCCTCCGAGGGGCTGGCTCGATGT 407
Db 3192 CTCCAGGAGAGAGGAGGCGCCAGCTCCACTGGCACCCTCCGAGGGGCTGGCTCGATGT 3251

Qy 408 ATTTGATGGTGACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAGAGGCTCCCCACACATGA 467
Db 3252 GTTTGATGGTGACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAGAGGCTCCCTCACAGA 3311

Qy 468 CCCCAGCCCTTACAGCGGTACAGTGAGAGACCCACAGTACCCCTGCCCTCTGAGACTGA 527
Db 3312 CCCCAGCCCTTCCACGCGGTACAGTGAGAGACCCCTACCGTACCCCTGCCCTCTGAGACTGA 3371

Qy 528 TGGCTACCTTGGCCCTTGGCTGACCTGACCTGACCTGACCTGAATATGTGAACCCAGATGT 587
Db 3372 TGGTAAGGTTGGCCCTTGGCTGACCTGACCTGACCTGACCTGAATATGTGAACCCAGATGT 3431

Qy 588 TCGGCCCCAGCCCTTGGCCCGGAGAGGGCCCTCTGCTGCTGCTGCCCGACCTGCTGGTGC 647
Db 3432 TGGCGCGCAGCCCTTGGCCCTAGAGGGCTTGGCTCTTCCCGACCGCTGGTGC 3491

Qy 648 CACTCTGGAAGG-----CCCAAGACTCTCTCTCCCGAGGAAGAATGGGT 692
Db 3492 CACTCTGGAAGGCGCCAAAGACTCTCTCTCCCGAGGAAGAATGGGT 3551

Qy 693 CGTCAAGAGCGTTTGGCTTGGGGGTGCGCTGGAGAACCCCGAG-ACCTTACACCCCA 751
Db 3552 TGTCAAGAGCGTTTGGCTTGGGGGTGCTGTGGAGAAATCCGGAGTACTTGGACCCCG 3611

Qy 752 GGGAGGAGTGCCTTCAGCCCGCCACCTCTCTCTGCTGCTTACGCCAGCGCTTGCACAACCT 811
Db 3612 GGGAGAGTGCCTTCAGCCCGCCACCTCTCTCTGCTGCTTACGCCAGCGCTTGCACAACCT 3671

Qy 812 CTATTACTGGGACAGGACCCAGAGAGGGGGGCTCCACCCAGCAGCTTCAAGAGGAC 871
Db 3672 GTATTACTGGGACAGGATCCATCAGAGCGGGGCTCTCCACCCAGCAGCTTGAAGGAC 3731

Qy 872 ACCTACGCGCAGAAACCCAGAGTACTTGGGTCTGGACGTGCCAGTGTGA 920
Db 3732 CCTACAGCAGAAACCCGGAGTACTTGGGGTGGACGTGCCAGTGTGA 3780

RESULT 10
AX189662
LOCUS AX189662 3771 bp DNA PAT 08-AUG-2001
DEFINITION Sequence 1 from Patent WO0148205.
ACCESSION AX189662
VERSION AX189662.1 GI:15143051
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 3771)
AUTHORS Spies, A.G.
TITLE Murine neu sequences and methods of use therefor
JOURNAL Patent: WO 0148205-A 1 05-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
1. 3771
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 808 a 1105 c 1068 g 790 t
ORIGIN

Query Match 23.9%; Score 508.8; DB 6; Length 3771;
Best Local Similarity 82.3%; Pred. No. 1.2e-115;
Matches 596; Conservative 0; Mismatches 127; Indels 1; Gaps 1;

Qy 198 GGGGTGGTCTTTTGGGATCTCTCATCAAGCGACGCGGAGAGATCAGATGTCAGACCC 257
Db 3048 GGAGTGGTGGTCTTGAAGAGTACTTGGTACCCCGCAGCAGGATCTTCTCCCGACCC 3107

Qy 258 TGCCCGGGCGCTGGGGCATGGTCCACACAGGACCCCGCAGCTCATCTACAGAGTGG 317
Db 3108 TGCCCGTAGGTACTTGGGAGCAGACCCCGCAGCAGCAGCTCTGTCGGCGCAGGAGTGG 3167

Qy 318 CGTGGGACCTGACATAGGCTGGAGCCCTTGAAGAGGAGGCCCGCAGCTTCTCCACT 377
Db 3168 CGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3227

Qy 378 GGCACCCCTCCGAAGGGCTGGCTCCGATGTATTTGATGGTACCTGGGAATGGGGSCAGC 437
Db 3228 GGCTCCCTCCGAAGGGCTGGCTCCGATGTATTTGATGGTACCTGGGAGTGGGGTAAC 3287

Qy 438 CAAAGGGCTGCAAGGCTCCCGCAGCAGATGACCCCGCTCTACAGCGGTACAGTGA 497
Db 3288 CAAAGGACTGCAGAGCCTCTCTCCACATGACCTCAGCCCTCTACAGCGGTACAGTGA 3347

QY 498 CCCACAGTACCCCTGCTGAGACTGATGGCTACGTTGCCCTGACCTGACGCC 557
|||||
Db 3348 TCCCAATACCTCTGCCCTGAGACTGATGGCTACGTTGCCCTGAGGCC 3407
QY 558 CCAGCCTGAATATGTGAACACCGCAGATGTTCCGGCCCGACGCCCTCGCCCGAGAGG 617
|||||
Db 3408 CCAGCCCGAGTATGTGAACCCAGAGGTTCCGGCTCAGTCTCCCTTGACCCCGAGAGG 3467
QY 618 CCCTCTGCTGCTGCCGACCTGCTGGTGCCACTGTGAAAGCCCAAGACTCTCTCCCC 677
|||||
Db 3468 TCCTCGCCTCCATCCGACCTGCTGGTGTCTACTAGAAAGACCAAGACTCTCTCTCC 3527
QY 678 AGGGAAGATGGGTCGTCGAAGAGCTTTTTCCTTTGGGGTGGCGTGAGAACCCCG- 736
Db 3528 TGGGAAAATGGGGTGTCAAGAGAGCTTTTTCCTTTGGGGTGGCTGGAGAACCTGA 3587
QY 737 AGACTTGACACCCAGGAGGAGCTGCCCTCAGCCACCCACCTCTCTGCTTCCAGGCC 796
|||||
Db 3588 ATACCTAGCACCCAGCAGGACTGCTCTCAGCCCAACCTCTCTGCTTCCAGGCC 3647
QY 797 AGCCTTGACAACTCTATTACTGGGACAGGACCCACAGAGCGGGGGTCCACCCAG 856
Db 3648 AGCCTTGACAACTCTATTACTGGGACAGAACTCATCGGAGCAGGTCTCCACCAAG 3707
QY 857 CACCTTCAAGGACACCTACGGCAGAGAACCCAGAGTACCTGGTCTGGAGTGGCCAGT 916
Db 3708 TACCTTTGAAGGACCCCACTGCGAGAAACCTGAGTACCTAGGCTGGATGTGCCAGT 3767
QY 917 GTGA 920
|||||
Db 3768 ATGA 3771

RESULT 11
LOCUS I21129 3955 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 14 from patent US 5518885.
ACCESSION I21129
VERSION I21129.1 GI:1601483
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3955)
AUTHORS Raziuddin;.. and Sarkar,F.H.
TITLE ERBB2 promoter binding protein in neoplastic disease
JOURNAL Patent: US 5518885-A 14 21-MAY-1996;
FEATURES Location/Qualifiers
source 1..3955
BASE COUNT 842 a 1147 c 1136 g 830 t
ORIGIN

Query Match 23.3%; Score 495.8; DB 6; Length 3955;
Best Local Similarity 83.0%; Pred. No. 2.1e-112;
Matches 57; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 232 CAGCAGAGATCACATGTCCAGACCCCTGCCCGGGCTGGGGGCTGGTCCACACAGG 291
|||||
Db 3110 CAGCAGGAGATTCTTCCCGGACCTTACCCAGGACTGGGAGCACAGCCATAGAAG 3169
QY 292 CACCGCAGCTCATCTACAGAGTGGCGTGGGACCTGACACTAGGCGTGGAGCCCTCT 351
|||||
Db 3170 CACCGCAGCTCGTCCACAGAGTGGAGGTGGTGAAGTGAAGTGGAGCCCTCG 3229
QY 352 GAAGAGGAGGCCCTCAGCTCTCCACTGCGACCTCCGAAGGGCTGGCTCCGATGATTT 411
|||||
Db 3230 GAAGAGGGCCCCCAGATCTCCACTGGCTCCCTCGGAAGGGGTGGCTCCGATGATTT 3289
QY 412 GATGGTGAACCTGGGAATGGGGGACCCAGAGGGCTGCAAAAGCCCTCCACACATGACCC 471
|||||
Db 3290 GATGGTGAACCTGGGAATGGGGGTAACCAAGAGGGCTGAGAGCCCTCTCCACATGACCTC 3349

QY 472 AGCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGACTGATGGC 531
|||||
Db 3350 AGCCTCTACAGCGGTACAGCGAGGACCCACATTACCTCTGCCCGGAGACTGATGGC 3409
QY 532 TAGTGTGCCCCCTGACCTGACGCCCCCAGCCTGAATATGTGAACCCAGAGATGTTGG 591
|||||
Db 3410 TATGTTGCTCCCTGGCTGCGAGCCCGAGCCCGAGTATGTGAACCAATCAGAGTTTCAG 3469
QY 592 CCCAGCCCCCTTCCGCCGAGAGAGGCCCTGCTGCTGCTGCCCGACCTGCTGGTGGCCACT 651
|||||
Db 3470 CCTGAGCCTCTTTAAACCCAGAGGGTCTCTGCTGCTGCTGCCGGCTGCTGGTGGTACT 3529
QY 652 CTGAAAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGGTCGTCAAAGACGTTTTTGCC 711
Db 3530 CTAGAAAGACCCAGACTCTCTCTCTCTGGGAAGATGGGGTGTCAAAGACGTTTTTGCC 3589
QY 712 TTTGGGGTGGCTGGAGAACCCCG-AGACTTGACACCCAGGAGGAGGAGTGGCCCCCTAG 770
Db 3590 TTCGGGGTGGCTGGAGAACCCCTGAATACTTAGTACCGAGAGAAGGCACTGCTCTCCG 3649
QY 771 CCCACACCTCTCTGCTGCTTCCAGCCAGCCTTCAAAGGGACACCTAGCGGACGAGAACCA 830
Db 3650 CCCACACCTCTCTGCTGCTTCCAGCCAGCCTTCAAAGGGACACCTAGCGGACGAGAAC 3709
QY 831 CCACAGAGCGGGGGCTTCCACCCAGCAGCCTTCAAAGGGACACCTAGCGGACGAGAACCA 890
Db 3710 TCATCGGAGCAGGGGCTTCCACCAAGTAACTTTGAAGGGACCCCACTGCGAGAACCT 3769
QY 891 GAGTACCTGGCTGGAGTGGCCAGTGTGAAGCCT 925
|||||
Db 3770 GAGTACCTGGCTGGATGTACCTGTATGAGACGT 3804

RESULT 12
LOCUS I59750 3955 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 14 from patent US 5654406.
ACCESSION I59750
VERSION I59750.1 GI:2478382
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3955)
AUTHORS Raziuddin;.. and Sarkar,F.Hoque.
TITLE Antibody to ERBB2 promoter binding factor
JOURNAL Patent: US 5654406-A 14 05-AUG-1997;
FEATURES Location/Qualifiers
source 1..3955
BASE COUNT 842 a 1147 c 1136 g 830 t
ORIGIN

Query Match 23.3%; Score 495.8; DB 6; Length 3955;
Best Local Similarity 83.0%; Pred. No. 2.1e-112;
Matches 57; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 232 CAGCAGAGATCACATGTCCAGACCCCTGCCCGGGCTGGGGGCTGGTCCACACAGG 291
Db 3110 CAGCAGGAGATTCTTCCCGGACCTTACCCAGGACTGGGAGCACAGCCATAGAAG 3169
QY 292 CACCGCAGCTCATCTACAGAGTGGCGTGGGACCTGACACTAGGCGTGGAGCCCTCT 351
Db 3170 CACCGCAGCTCGTCCACAGAGTGGAGGTGGTGAAGTGAAGTGGAGCCCTCG 3229
QY 352 GAGAGAGGAGGCCCTCAGGCTCTCCACTGGCAGCCTCCGAAGGGCTGGCTCCGATGATTT 411
Db 3230 GAAGAGGGCCCCCAGATCTCCACTGGCTCCCTCGGAAGGGGTGGCTCCGATGATTT 3289
QY 412 GATGGTGAACCTGGGAATGGGGGACCCAGAGGGCTGCAAAAGCCCTCCACACATGACCC 471
|||||


```

QY 532 TAGTTGGCCCCCTGACCTGCAGCCCCCAGCGTGAATATGTGAACCCAGACGATGTTGCG 591
Db 3410 TATGTTGCTCCCTGGCTGCAGCCCCCAGCGTGAATATGTGAACCCAGACGATGTTGCG 3469
QY 592 CCCCAGCCCCCTTCCCGCGAGAGAGGCGCTCTGCTGCTGCTGCCGACCTGTGTCGCACT 651
Db 3470 CTTAGCCTCTTAAACCCAGAGAGGCTCTGCTGCTGCTGCCGACCTGTGTCGCACT 3529
QY 652 CTGGAAGGCCCAAGACTCTCTCCAGGGAAGATGGGTCTGCAAGAGCTTTTGGC 711
Db 3530 CTAGAAGACCAAGACTCTCTCTCTGGGAAGATGGGTCTGCAAGAGCTTTTGGC 3589
QY 712 TTTGGGGTGGCTGGAGAACCCCG-AGACTGTGACACCCAGGAGGAGCTGCCCTCAG 770
Db 3590 TTCGGGGTGGCTGGAGAACCCCTGTAATCTAGTACCGAGAGGACCTGCTCTCG 3649
QY 771 CCCACCCCTCTCTGCTTTCAGCCAGCCTTCGACACCTCTATTACTGGGGACCAAGAC 830
Db 3650 CCCACCCCTCTCTGCTTTCAGCCAGCCTTCGACACCTCTATTACTGGGGACCAAGAC 3709
QY 831 CCACCAGAGGGGGGCTCCACCAGACCTTCAAGGAGACCTACGGCAGAGAACCCA 890
Db 3710 TCATCGGAGAGGGGGCTCCACCAGAACTTGAAGGAGACCCCTACGAGAACCT 3769
QY 891 GAGTACCTGGCTTGGAGCGCCAGTGTGAAGCT 925
Db 3770 GAGTACCTAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3804

RESULT 14
LOCUS HAMNEU 4062 bp mRNA ROD 01-FEB-2000
DEFINITION Syrian golden hamster neu mRNA for p-185, complete cds.
ACCESSION D16295
VERSION D16295.1 GI:493236
KEYWORDS EGF-receptor; cellular oncogene; neu-differentiation
factor/hergulin; p-185; tyrosine kinase.
SOURCE Mesocricetus auratus (individual isolate animal 14) peripheral
nerve neoplastic Schwann cell cell-line 14-2 (library: NIH3T3) cDNA
to mRNA, clone pSHN.
ORGANISM Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
1 (bases 1 to 4062)
Nakamura,T., Ushijima,T., Ishizaka,Y., Nagao,M., Arai,M.,
Yamazaki,Y. and Ishikawa,T.
Cloning and activation of the Syrian hamster neu proto-oncogene
Gene 140 (2), 251-255 (1994)
94193007
2 (bases 1 to 4062)
Nakamura,T.
Direct Submission
Submitted (19-MAY-1993) to the DDBJ/EMBL/GenBank databases. Takuro
Nakamura, Faculty of Medicine, University of Tokyo, Department of
Pathology, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113, Japan
(Tel:03-3812-2111(ex.3356), Fax:03-3815-8379)
Submitted (19-May-1993) to DDBJ by:
Takuro Nakamura
Department of Pathology
Faculty of Medicine, University of Tokyo
7-3-1 Hongo
Bunkyo-ku Tokyo 113
Japan
Phone: 03-3812-2111 x3356
Fax: 03-3815-8379.
Location/Qualifiers
1. .4062
/organism="Mesocricetus auratus"
/isolate="animal 14"
/db_xref="taxon:10036"
/cell_line="14-2"
FEATURES
source

```

```

/cell_type="neoplastic Schwann cell"
/clone_lib="NIH3T3"
/tissue_type="peripheral nerve"
33..3797
/gene="neu"
33..86
/gene="neu"
/product="putative"
33..3797
/gene="neu"
/codon_start=1
/product="p-185 precursor"
/protein_id="BAA03801.1"
/db_xref="GI:747595"
/translation="MELAAWCGMLLALLSPGASGTQVCTGDMKRLPASPETHLD
IVRLYQGVQVGNLELYLPANATLTFLQDIOEVOGFMILIAHSHVRHVPQRLRIY
RGTFLEDKYALAVLDRPDLNVNATATGTGEGRELOLRSLTELKGVLLIRNPQ
LCYQDYLWKDVPRKNNQAPVDIDTNRSRPCPCAPACKDNHGWGSPEDCQTLGT
IAPRAVPAARARLPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTNT
DTFSPNPEGRYTFGASCTTCCPNYLSFVSGCTLVCLNNQVTAEDGTORCEK
SKSARVCYGLGMEHLRGARAITSANIOFAGCKIFGSLAFLPESFDGNSGGLPL
TPQLQVFTLEITGLYISAWPDSLHDLVSFQNLVIRGRVLRHDSYSLALQGLI
RWGLSLRLGSLVLIHRNTHLCFVHTVPMDLFRNPHQALLHSGNPSLEEGLKD
FACYPLCAHGHGPGPCTQVNCNSHFLRQBCVKECVKWLPLREYVNGKCLCPHE
COPONSTECTSEADOCCTACPHYKDSPPCARCSPGPDLSYMPIWKYPDERMCQ
PCPINCTSCVDLDERGCPAEORASPATSIATVVGILLFLVGVGILIKRRKRI
RAYTMRKLIQTELVEPLTPSGAPNQAQMRILKTELKRVKVLGSGAGFVYKGIWI
PDGENKIPVAIKVLENTSPKANKEILDEAYVMAGLSGVPYSLRGLCJLSTVOLY
OLMPYGCCLDHVREHGRGLSODLLNWCVOIAGMSYLEYLRVLRHDAARNLVKSP
NHVITDFGLARLDDIETEVHADGKVPKIKWIALESILRRRTHSDVNSIGVTVME
LMTFGAPYDGPAREIPDLLEKGRPLPOPICTIDVIMVKMMDISECRPREL
VSEFARMARDPQFVVIQNLDELSPSSPLDFTFYRLSLEDMDGLDAEYLVPQGF
PEPDAPGAGSTAHRRHSSTRSGGELTMEPEPGEPPRSPPLAFSEGAGSVFEG
ELMGATGPOSISPRDLSPIQRYSEDTLPETPTDGYAVPLACSPQEVVNOPEVR
SOPPLTPEGPUPPVPRPAGATLERPKTLSPKNGVVKDVFYFGAVENPEYLVPRGSA
SOPHPPLCAPFDNLYYWDQPSRSGSPNTFECTPTAENPEYLGLDVPV"
87..3794
/mat_peptide
/gene="neu"
/product="p-185"
516..1988
/gene="neu"
/feature="extracellular domain"
1989..2057
/gene="neu"
/feature="transmembrane domain"
2058..3794
/misc_feature
/gene="neu"
/feature="cytoplasmic domain"
BASE COUNT 845 a 1209 c 1182 g 826 t
ORIGIN

```

```

Query Match 22.9%; Score 487; DB 10; Length 4062;
Best Local Similarity 82.4%; Pred. No. 3.2e-110;
Matches 571; Conservative 0; Mismatches 120; Indels 2; Gaps 1;
QY 232 CAGCAGAGATCATGTCCAGACCCCTGCCCGGGCGCTGGGGCGCATGGTCCACACAG 291
Db 3111 CAGCAGGATTTCTTCTCCAGACCCCTGCCAGCGCGCGGACGCGCCACCGTAGG 3170
QY 292 CACCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTGACATAGGCTGGAGCCCTCT 351
Db 3171 CACCGTAGCTCGTCAACAGGAGTGGAGCGGCGGAGCTGACATGGGCGATGGAGCCCTCG 3230
QY 352 GAACAGAGGCCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGCTGGCTCCGATGATTT 411
Db 3231 GGAGAGAGCCCCCAGGTCTCCACTGGCTCCCTCGAAGGGCTGGCTGATGTTT 3290
QY 412 GATGCTGACCTGGAGTGGGCGAGCCAGGGGCTGCCAAGGCTCCCCACACATGACCCC 471
Db 3291 GAGGTGAAGTGGGATGGGCGTACCAAGGGCGCGAGCATCTCTCCAGGTGACCTC 3350
QY 472 AGCCCTCTACGCGGTACAGTGAGGAGCCCCACAGTACCCCTGCCCTCTGAGACTGATGCG 531

```



```

RESULT 15
AX027785/c
LOCUS
AX027785 11795 bp DNA
Sequence 4 from Patent WO0043527.
ACCESSION
AX027785
VERSION
AX027785.1 GI:10188633
KEYWORDS
synthetic construct.
SOURCE
synthetic construct.
ORGANISM
artificial sequence.
REFERENCE
1 (bases 1 to 11795)
AUTHORS
Bollen,A., Jacobs,P., Jacquet,A., Haumont,M. and Massaer,M.G.
TITLE
Varicella-zoster virus vaccines
JOURNAL
Patent: WO 0043527-A 4 27-JUL-2000;
SMITHKLINE BECHAM BIOLOG (BE) ; BOLLEN ALEX (BE) ; JACOBS PAUL
(BE) ; JACQUET ALAIN (BE) ; HAUMONT MICHELE (BE) ; MASSAER MARC
GEORGES FRANCIS (BE)
FEATURES
Location/Qualifiers
1..11795
/source="synthetic construct"
/db_xref="taxon:32630"
2917 a 2760 c 2957 g 3161 t
BASE COUNT
ORIGIN

```

	Query Match	21.4%	Score 454.8	DB 6	Length 11795	
	Best Local Similarity	99.6%	Pred. No. 2.8e-102			
	Matches 456	Conservative	0	Mismatches	0	Gaps 0
Qy	1411	GTGTAGCCGTTAGTGTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT	1470			
Db	986	GTGTAGCCGTTAGTGTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT	927			
Qy	1471	CTGCTAATCTCTGTTTACCAGTAGGCTGTGCCAGTGGCGATAAGTGTGCTTACCGGGTTG	1530			
Db	926	CTGCTAATCTCTGTTTACCAGTAGGCTGTGCCAGTGGCGATAAGTGTGCTTACCGGGTTG	867			
Qy	1531	GACTCAAGACGATAGTTTACCGGATAAGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTGC	1590			
Db	866	GACTCAAGACGATAGTTTACCGGATAAGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTGC	807			

Search completed: January 17, 2002, 11:49:53
Job time: 17205 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:03:06 ; Search time 599.86 Seconds
(without alignments)
3037.069 Million cell updates/sec

Title: US-09-242-202A-22

Perfect score: 2125

.Sequence: 1 GCCACCATGCCCTGACCT.....AGCCTCTCCACAGGTACCT 2125

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*

- 1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2125	100.0	2125	19 AAV21727	Humanised vector p
2	1203	56.6	1425	19 AAV21724	Humanised vector p
3	1194.4	56.2	1911	19 AAV21726	Humanised vector p
4	671.2	31.6	4299	14 AAQ46083	Sequence encoding
5	671.2	31.6	4473	20 AAZ31071	HER-2 nucleic acid
6	670.4	31.5	9274	22 AAF24297	HER2 transgene pla
7	669.6	31.5	4530	16 AAT01585	Her-2/neu (ERBB2)/c
8	669.6	31.5	4530	18 AAT71253	Human HER2 gene.
9	669.6	31.5	4530	21 AAZ60815	Nucleotide sequenc
10	669	31.5	4472	21 AAA14812	cDNA encoding the
11	667.6	31.4	3768	17 AAT40739	HER-2/neu oncogene

12	667.6	31.4	3768	20 AAX01912	Human HER-2/neu on
13	667.6	31.4	3768	22 AAH23392	Human HER-2/neu pr
14	666	31.3	3768	21 AAA09455	Human heregulin 2
15	508.8	23.9	3771	21 AAA89737	Mouse Her-2/neu CD
16	508.8	23.9	3771	22 AAH42210	Nucleotide sequenc
17	499.6	23.5	3600	21 AAA89736	Human HER-2/neu co
18	495.8	23.3	3955	16 AAT01590	Rat neu promoter.
19	495.8	23.3	3955	21 AAA89753	Rat HER-2/neu prot
20	474.8	22.3	1115	21 AAF21778	Human breast and o
21	459.8	21.6	3553	15 AAQ74208	Expression vector
22	454.8	21.4	7073	18 AAT50962	TF8-SG9 CDR-grafte
23	454.8	21.4	7864	18 AAT50963	TF8-SG9 CDR-grafte
24	454.8	21.4	11795	21 AAA59345	Nucleotide sequenc
25	454.8	21.4	13254	17 AAT40915	Nucleotide sequenc
26	454.8	21.4	13254	21 AAA32151	pBel2 Combo BM 12
27	454.8	21.4	13254	21 AAA32165	Complete nucleotid
28	453.8	21.4	608	19 AAV21731	ColE1 origin of re
29	453.8	21.4	5241	17 AAT27307	pHL104 containing
30	453.8	21.4	5365	10 AAN90646	Nucleotide sequenc
31	453.8	21.4	5413	10 AAN90649	Nucleotide sequenc
32	453.8	21.4	5518	10 AAN90647	Nucleotide sequenc
33	453.8	21.4	5566	10 AAN90648	Nucleotide sequenc
34	453.8	21.4	5579	11 AAQ02032	Plasmid pHS 164.
35	453.8	21.4	6051	12 AAQ14934	pBG391 sequence.
36	453.8	21.4	6149	10 AAN90645	Sequence of T4 lym
37	453.8	21.4	6151	10 AAN90644	Sequence of T4 lym
38	453.8	21.4	6151	11 AAQ03006	Full length T4 CDN
39	453.8	21.4	6151	11 AAQ05608	Plasmid pBG381 inc
40	453.8	21.4	6165	13 AAQ20324	Sol.rhesus CD4 seq
41	453.8	21.4	6477	11 AAQ02030	Plasmid pinf 4-49.
42	453.8	21.4	7252	20 AAX86928	Complete sequence
43	453.8	21.4	7377	11 AAQ03005	Full length T4 CDN
44	453.8	21.4	7377	11 AAQ05607	Plasmid p170-2 inc
45	453.8	21.4	8119	17 AAT33477	Subtilisin N62D/G1

ALIGNMENTS

RESULT 1

AAV21727

ID AAV21727 standard; cDNA; 2125 BP.

XX AC AAV21727;

XX DT 17-AUG-1998 (first entry)

XX DE Humanised vector pITL-hHER/neu.

XX KW Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;

XX KW human; HER-2/neu; C-erbB-2; breast cancer; ds.

XX OS Chimeric - Homo sapiens.

XX OS Chimeric - Escherichia coli.

XX OS Synthetic.

XX FH Key

XX CDS Location/Qualifiers

XX FT 13..921

XX FT /*tag= a

XX FT /product= human HER-2/neu

XX FT 922..1181

XX FT /*tag= b

XX FT /note= "combined splice and polyA sequences"

XX FT 1195..1401

XX FT /*tag= c

XX FT /note= "SupF gene"

XX FT misc_feature

XX FT 1412..1864

XX FT /*tag= d

XX FT /note= "ColE1 origin of replication"

XX FT 1877..2125

XX FT /*tag= e

XX FT /note= "RANTES promoter"

PN W09806863-A1.
XX 19-FEB-1998.
XX 14-AUG-1997; 97WO-US14306.
XX 14-AUG-1996; 96US-0023931.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX Example 7; Page 41-42; 125pp; English.
XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
XX a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
XX evaluate the toxicity of anti-tumour vaccination in rats, and in
XX phase I and phase II trials to evaluate polynucleotide vaccination
XX in advanced breast cancer. Novel humanised vectors, which can be
XX based on pITL, comprise a human-derived promoter or mammalian
XX homologue which is functional in mammalian target tissue and cells
XX and a sequence acceptance site which accepts cDNA products from
XX RT-PCR cloning. The vectors are non-replicating in mammalian cells
XX but are capable of extended stable expression of the target
XX sequence, generating an immune response in immunised individuals.
XX The vectors selectively elicit immune responses to the target
XX sequences with little or no immune response to the other components
XX of the vectors.
XX Sequence 2125 BP; 449 A; 550 C; 598 G; 428 T; 0 other;

Query Match 100.0%; Score 2125; DB 19; Length 2125;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACATGGCCCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGAGGGC 60
Db 1 gccaccatggccctgaacctctctacatgccatctggaagtgtccagatgagagggc 60

Qy 61 GCATGCCAGCTTGCCCATCACTGACCCACACCTCTGTGGACCTGGATGACAAAGGC 120
Db 61 gcatgcccagcttgcccatcaactgacccactcctgtgtggacctggaagagggc 120

Qy 121 TGCCCGCCGACGAGACGCCCTCTGACGTCCATCATCTCTGCGGTGGTGGCAAT 180
Db 121 tgccccgacgagagagccacctgacgtccatcatctctcggtgtggcaat 180

Qy 181 CTGCTGGTGTGCTGTGGGGTGGTCTTTTGGGATCTTCATCAAGCGACGCGAGAG 240
Db 181 ctgctggtgtgctgtgggtgtgtcttggatccctcatcaagcgagcgaggaag 240

Qy 241 ATCATGTCCAGACCTGCCCCGGCGCTGGGGCATGCTCCACACAGCGACCGGAGC 300
Db 241 atcatgtccagacctgccccggcgctggggcatgctccacacagcgacccgagc 300

Qy 301 TCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 360
Db 301 tcactaccagagtggcgtgaggacctgacactagagctgagccctctgaagaggag 360

Qy 361 GCCCGCAGGTCTCCATGGGACACCTCCGAAGGGGTGGCTCCGATGATTTGATGGTAC 420
Db 361 gccccaggtctccactggcaccctccgaagggtgtgctccgatgtatttgatggtgac 420

Qy 421 CTGGAATGGGGGAGCAAGGGCTGCAAGAGCTCCCGACACATGACCCGAGCCCTCTA 480
Db 421 ctgggaatgggggagccaagggtgtgcaaggctccccacacatgaccccgacctcta 480

Qy 481 CAGCGTACAGTGAAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGTACGTTGCC 540
Db 481 cagcgtacagtgaagacccacagtagtacctgacctgacctgacctgacctgacctg 540

Qy 541 CCCTGACCTGCAGCCCGCCAGCCTGAATATGTGAACAGCCAGATGTTCCGCCCGAGCC 600
Db 541 ccctgacctgagccccagcctgaatgtgaaccagccagatgttcggccccagccc 600

Qy 601 CTTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGTGGTCCACCTCTGAAAGG 660
Db 601 ctttcgccccagagggccctctgctgctgccccagacctgctgctgcaactctgaaag 660

Qy 661 CCCAAGACTCTCTCCCGAGGGAAGAAATGGGTGCTGCTAAAGAGTGTTCCTTTGGGGGT 720
Db 661 cccaagactctctcccgagggaagaatgggtcgctcaagacgtttttgctttgggggt 720

Qy 721 GCCGTGAGAACCCCGAGACTTGACACCCAGGAGGAGCTGCCCTCAGCCGCCACCCCTC 780
Db 721 gccgtgagaaccccgagacttgacacccagggagagctgccccctcagccccacctc 780

Qy 781 CTCCTGCTTCAGCCCGCCAGCCTTCGACACCTCTATTACTGGACCGACGCCACACAGC 840
Db 781 ctcctgcttcagccccagccttcgacaacctattactggaccagaccaccagagc 840

Qy 841 GGGGGGTCTCCACCGACACTTCAAGGGACACCTACGGCAGAGAACCCAGTACTGG 900
Db 841 ggggggtctccacccagaccttcaaggagcacctacgagcagagaaacccagagtaactgg 900

Qy 901 GTCTGGAGTCCCGAGCTGAAGCCTTAAGGCCATATGCTGAGTGGATGCTTGACCCCA 960
Db 901 gtctggagctcccgagctgaagccttaaggccatatggtgagtggaagccttgacccca 960

Qy 961 GGCGGGATGGGGAGACCTGTAGTCAGAGCCCGCGGACGACACAGGCCAATGCCCGTCC 1020
Db 961 ggcgggatggggagacctgtagtcagagccccggcgagcacagagccaatgccgtcc 1020

Qy 1021 TTCCCTGCACTGAGTACTGCTGCCGGGTGGATCCTGTGACCCCTCCCGAGTGCT 1080
Db 1021 ttccctgcactgagtagtactgctgccgggtgggatccctgtgacccctcccagtgct 1080

Qy 1081 CTCCTGCGCTGGAGTTGCCACTCCAGTGCCACCGCTTCTCTAATAAATAAATTAAGT 1140
Db 1081 ctctgctgctggagttgcccactccagtgcccaccagcctgtctctaataaataaagt 1140

Qy 1141 TGCATCATTTTGTCTGACTAGGTGCTCTATAATATATTAAGCTTGATATCGAATCTT 1200
Db 1141 tgcatacttttgtctgactaggtgctctataataattataagcttgatatacgaattctt 1200

Qy 1201 TCTCAACGTAACTTTACAGGGCGCTCATTTGATATGATGCGCCCGCTTCCCGATA 1260
Db 1201 tctcaagtaaacctttacagcgcgctcatttgatatgatgagcgcccgcttcccgata 1260

Qy 1261 AGGAGCAGGCGCAGTAAAGCATTAACCGTGGGTTCGCGAGCGCCGAAAGGAGCA 1320
Db 1261 aggagcagcgccagtaaaagcattaccctggtgggttccccgagcgcccaagagca 1320

Qy 1321 GACTCTAAATCTGCGCTCATCGACTTCGAAGGTTGGAATCTTCCCGCACCCATCACT 1380
Db 1321 gactctaaatctgctgactgacttcgaagtttgaatccttccccaccaccatcaact 1380

Qy 1381 TTCAAAAGTCGAAGAATTCCTGAGCCCGTGTAGCCGTAGTAGGCCACCATCTCAAG 1440
Db 1381 ttcaaaagtcgaaagaattctcgagcccgtagcgctgtagcgctgtaggccaactcaag 1440

Qy 1441 AACCTGTAGCACCGCTACATACCTGCTCTGCTTAATCTGTTACCATGCTGCTGCC 1500
Db 1441 aacctgtagcacccgctacatacctgctctgctaaacctgtaccagtgctgctgcc 1500

Qy 1501 AGTGCGATAAGTCTGTCTTACCGGGTTGACTCAAGACGATAGTACCAGTAAGGCG 1560
Db 1501 agtgcgataagtcgtgtcttaccgggttgactcaagacgagatgattaccggaagagcg 1560

Qy 1561 CAGCGTGGGCTGAACGGGGGTTCTGTGCACACAGCCCGCTTGGAGCGAACGACCTAC 1620

```
|||||
Db 1561 cagcggctcgggctgaacggggggttcgtagcacacagccagcttgagcgaacgacctac 1620
Qy 1621 ACCGAACGTAGATACCTACAGCGTGAGCATTTGAGAAAGCGCCACCGCTTCCCGAAGGGAGA 1680
Db 1621 accgaactgagatacctacagcgtgagcattgagaaagcgcacgcttcccgaaggagga 1680
Qy 1681 AAGCGGACAGGTATCCGTAAGCGGCGAGGTCGCAACAGGAGAGCGCACGAGGAGGCTT 1740
Db 1681 aagcggacaggtataccggtaagcggcagggttcgaaacagggagcgcacagggagcctt 1740
Qy 1741 CCAGGGGAAACCGCTGGTATCTTTATAGTCTCTGTCGGGTTTTCGCCACCTCTGACTTGAG 1800
Db 1741 ccaggggaaacgcgctgatactttatagtcctgctcggttttcgccacctctgacttgag 1800
Qy 1801 CGTCATTTTGTGATGCTGTCGAGGGGGGCGGAGCCCTATGGAATAACGCCAGCAACGGG 1860
Db 1801 cgtcattttgtgatgctcgtcagggggcgagcctatggaaaaacgcgcaacgcgcg 1860
Qy 1861 GCCGGGATCCGGAGAGTCACTCTAGATGAGAGAGCAGTGAGGAGAGACAGAGACTC 1920
Db 1861 gccggggatccggagagcctcactctagatgagagcagtgagggagagacagagactc 1920
Qy 1921 GAATTTCCGGAGCTATTTCACTTTTCTTTCCGTTTGTGCAATTTCACTTATGATACCG 1980
Db 1921 gaatttcggagactattcagtttcttcttccgtttgtgcaatttcacttatgataccg 1980
Qy 1981 GCCATGCTGGTGGCTATTTTGGAAACTCCCTTATAGGGGATGCCCTTCACTTATGCGCCTA 2040
Db 1981 gcaaatgctgtgtgctattttggaactcccccttaggggatgcccccaactgagcccta 2040
Qy 2041 TAAAGGGCCAGCTCAGCTGCGAGAGGATTCCTGCGAGAGATCAAGACAGCAGCTGGACCT 2100
Db 2041 taaaggccagcgtgagctgcagagagattccctgcagagatcaagacagcagctggacct 2100
Qy 2101 GCGACAGCTCTCCACAGGTACCT 2125
Db 2101 cgcacagcctctccacaggtacct 2125
```

RESULT 2

AAV21724
ID AAV21724 standard; cDNA; 1425 BP.

AC AAV21724;

XX 17-AUG-1998 (first entry)

XX Humanised vector pITL.

XX Vector; vaccine; tumour; antigen; plasmid pITL; ds.

XX Chimeric - Homo sapiens.

OS Chimeric - Escherichia coli.

OS Synthetic.

FH Key Location/Qualifiers

FT misc_feature 1..221

FT /*tag= a

FT /note= "stuffer sequence"

FT 22..481

FT /*tag= b

FT /note= "combined splice and polyA sequences"

FT 495..701

FT /*tag= c

FT /note= "SupF gene"

FT 712..1164

FT /*tag= d

FT /note= "ColE1 origin of replication"

FT 1177..1425

FT /*tag= e

FT /note= "RANTES promoter"

PN WO9806863-A1.
XX 19-FEB-1998.
PD 14-AUG-1997; 97WO-US14306.
XX 14-AUG-1996; 96US-0023931.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Nelson EL, Nelson PJ;
PI WPI; 1998-159552/14.
DR Humanised polynucleotide vectors - comprising human derived promoter
PT and sequence acceptance site, used for the production of vaccines
XX Claim 14; Page 29-30; 125pp; English.
XX Plasmid pITL comprises a base vector for novel humanised
CC polynucleotide vectors. Such vectors comprise a human-derived
CC promoter or mammalian homologue which is functional in mammalian
CC target tissue and cells and a sequence acceptance site (see
CC AAV21735-36), which accepts cDNA products from RT-PCR cloning. They
CC also contain minimal non-human components, such as a replication
CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
CC are necessary for production of the vector, as well as human-derived
CC splice and polyA sequences (see AAV21723). The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX Sequence 1425 BP; 325 A; 381 C; 404 G; 315 T; 0 other;

Query Match

Best Local Similarity 56.6%; Score 1203; DB 19; Length 1425;

Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 922 GCCTTAAGGGCCATATGTTGAGTGGATGCTTGACCCCGAGGGGATGGGGAGACCTG 981

Db 223 gcttaaggggccatatggtgagtgagcttgaccccgaggggagagacctg 282

Qy 982 TAGTCAGAGCCCCCGGCGACGACAGCCCAATGCCCTTCCCTGCAGTGTAGTGA 1041

Db 283 tagtcagagccccggggcagcagagcccaatgccctccctccctgcagtgagtagta 342

Qy 1042 CTGCCCGGGTGGGATCCCTGTGACCCCTCCCGAGTGCCTCTCTGTGCCCTGGGAAGTTGCC 1101

Db 343 ctgcccggtgggatccctgtgacccctccctgcctctcctcctgcctgcctgcctg 402

Qy 1102 ACTCCAGTGCACACAGCCTTGTCTTAATAAATTAAAGTTGCATCATTTTGTCTGACTAG 1161

Db 403 actccagtgccccaccagcctgtcctaataaaataaagttgcatcatcttctgtgactag 462

Qy 1162 GTGTCCTCTATATATATAAGCTTGCATATCGAATTTCTTCAAGTAACTTTACAG 1221

Db 463 gtgtcctctataataataaagcttgatcgaatctcttctcaacgtaacactttacag 522

Qy 1222 CGGCGCGTCATTTGATATGATGCGCCCGCTTCCCGATAGGGAGCAGGCGCAATAAAGC 1281

Db 523 cggcgctcatttgatgatgatgcgcccgcttcccgataaaggagcagcagcagtaaaagc 582

Qy 1282 ATTACCCGTTGGTGGTTCGCCAGCGGCCAAAGGGAGCAGACTCTTAATGTCCGTCATC 1341

Db 583 attacccgtggtgggttcccgagcgccaaaggagcagactcctaatactctgcgtcatc 642

QY	1342	GACTTCGAAGTTTCGATCCTTTCCCCACACACCATCACTTTCAAAAGTCGGAAGAATTC	1401
Db	643	gacttcgaagtttcgaatccttccccccaccatcactttcaaaagtcggaagaattc	702
QY	1402	CTGCAGCCGTGTAGCCGTAGTTAGCGCACCACTCAAGAACTCTGTAGCACCGCCTACA	1461
Db	703	ctgcagccggttagccgttagttagccaccacttcaagaactctgtagcacgcgctaca	762
QY	1462	TACCTCGCTCTGCTAATCCTGTTTACAGTGGCTGTCGCACGTGGCGATAAGTCTGTGCTT	1521
Db	763	tacctcgctctgctaactctgttaccagtggctgctgccagtggcgataagtcgtgtctt	822
QY	1522	ACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGCGCACAGCGTCGGCTGAACGGGG	1581
Db	823	accgggttgactcaagacgatagtttacgcgataagcgcgagcgtcgggctgaacgggg	882
QY	1582	GGTTCTGTGCACACAGCCGACGTTTGAGGCGAGGACCTTACACCGAACTGAGNACTACAG	1641
Db	883	ggttctgtcacacagcccagcttggagcgaaacctacacccgaactgagataacctacag	942
QY	1642	CSFTGACATTGAGAAAGCCGACGCTTCCGAAGGAGGAAAGCGGACAGGTATCCGGTA	1701
Db	943	cgtgacatttgaaaagcgccacgcttccccaaagggaagaaaggcgacaggtatccggta	1002
QY	1702	AGCGGACGGTTCGAAACAGAGAGCGACGAGGGAGCTTCCAGGGGAAACCGCTGGTAT	1761
Db	1003	agcggcagggttcgaaacagagagcgacagaggagcttccagggggaaacgcctggtat	1062
QY	1762	CTTTATAGTCTCTCGGGTTTCGCCACCTCTGACTTGAGCTGCTGATTTTGTGATGCTCG	1821
Db	1063	ctttatagtctctcggttctgccacctctgacttgagctcgatcttttgtgatgctcg	1122
QY	1822	TCAGGGGGCGGAGCTATGGAANAACCCACACGCGCGCGGGGATCCGGAGAGCTC	1881
Db	1123	tcaggggggcgagcctatggaaaacgccaacgcgcccggggatccggagagctc	1182
QY	1882	ACTCTAGATGAGAGACGCTGAGGGAGAGACAGACTCGAATTTCCGGAGCTATTTTCAG	1941
Db	1183	actctagatgagagagcagtggaggagagacagagactcgaatttcggagctatttcag	1242
QY	1942	TTTCTTTTCCGTTTGTGCAATTTACATTATGATACCGGCCAATGCTTGTTGCTATTTT	2001
Db	1243	ttttctttccgttttgtgcaatttcaacttatgataccgggccaatgcttgggttgctattt	1302
QY	2002	TGGAAATCCCTTAGGGATGCCCTCAACTGGCCCTATTAAGGGCCAGCCTGAGCTGC	2061
Db	1303	tggaaatccctctagggtatgcccctcaactggccctataagagggccagcctgagctgc	1362
QY	2062	AGAGGATTCTGCAGAGGATCAAGACAGCAGCTGGACCTCGCAGAGCTCTCCACACAGT	2121
Db	1363	agagattcttcagagatcaagacagacagcagtgaacctcgcacagcctctcccaaggt	1422
QY	2122	ACC 2124	
Db	1423	acc 1425	
RESULT 3			
AAV21726			
XX	ID	AAV21726 standard; cDNA; 1911 BP.	
XX	AC	AAV21726;	
XX	XX		
DT	17-AUG-1998	(first entry)	
XX	DE	Humanised vector pITL-GFP.	
XX	DE		
KW	KW	Vector; vaccine; tumour; antigen; plasmid pITL-GFP;	
XX	XX	green fluorescent protein; ds.	
OS	OS	Chimeric - Homo sapiens.	
OS	OS	Chimeric - Escherichia coli.	
OS	OS	Chimeric - Aequorea victoria.	

[illegible]

12

||||| 836 ggtggatccctgtgaccctccctgctctctgctgacctggaagtgcacctccag 895
QY 1109 TGCCACCAGCCTTGCCTAAATAAATTAAGTTGCATCATTTTGTCTGACTAGGTGCTCT 1168
Db 896 tgcaccagcctgttcccaataaaataagttgcatcattttctgactaggtgtcct 955
QY 1169 CTATAATATATAAGCTTGATATCGAATCTTCTCAACGTAACTTTACAGCGCGCG 1228
Db 956 ctataattataagcttgatctcgaattcttctcaacgttaacactttacagcgcg 1015
QY 1229 TCATTTGATATGATGCGCCCGCTTCCGATAAAGGAGCAGCCAGTAAAGCATTTACCC 1288
Db 1016 tcatttgatgatggtccccgttccgataaggagcagcgcaataaagcattaccc 1075
QY 1289 GTGGTGGGTTCGAGCGCGCAAGGAGCAGACTCTTAATCTGCGTCATCGACTTCG 1348
Db 1076 gtggtggtgtcccgagcgccaaaggagcagactctaaatctgcgtcatcgactcg 1135
QY 1349 AAGTTTCGAATCTTCCCCACCACCATCTTCAAAAGTCCGAAAGAAATTCCTGCAGC 1408
Db 1136 aaggttcgaatcttccccaccaccatcaacttcaaaagtcgaaagaattcctgcagc 1195
QY 1409 CCGTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTTACATACCTCG 1468
Db 1196 ccgtgtagccgttagggccaccacttcaagaactctgtagcaccgctacatacctcg 1255
QY 1469 CTCTGTATCTCTTTACAGTGGCTGCTGCCAGTGGCGATAAGTGTCTTTACCGGGT 1528
Db 1256 ctctgtatctctgttaccagtggtcgtcgcaagtgcgataaagtcgtcttaccgggt 1315
QY 1529 TGGACTCAAGACATAGTTACCGGATAAAGCGCAGCGGTGCGGGCTGAACGGGGGTTCGT 1588
Db 1316 tgaactcaagacagatgttaccgataaagcgagcggtcggtcaagcggttctgt 1375
QY 1589 GCACAGCGCCACTTGGAGCGAAGCACTTACACCGAACTGTAGATACCTTACACGCTGAGC 1648
Db 1376 gcacacagccagcttggagcgagcagcctacaccgaaactgagatacctacagcgtgagc 1435
QY 1649 ATTGAGAAAGCGCCAGCTTCCGAGAGGAGAAAGCGGACAGGTATCCGTTAAGCGGCA 1708
Db 1436 attgaaagcgccacgcttcccgagggagaaagcgagaggtatccggttaagcgca 1495
QY 1709 GGTCGGAACAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATA 1768
Db 1496 ggtcgaaacagagagcgacagagaggttccagggggaaagcgtggtatcttata 1555
QY 1769 GTCCGTGCGGTTTCCGACCTCTGACTTGAGCGTGTGATTTTGTGATGCTGTACAGGG 1828
Db 1556 gtctgtcggtttcgccacctgtgactgagcgtcgatattttgtgctgctcagggg 1615
QY 1829 GCGGAGCCTATGGAAGAGCCAGACCGCGCGCGGGGATCCGGAGAGCTCACCTCTAG 1888
Db 1616 ggagcgctatggaagacgccaagcagcgcgcggggagatccgagagctcactctag 1675
QY 1889 ATGAGAGAGCAGTGTAGGGAGAGACAGAGACTCGAAATTTCCGGAGCTATTTCAGTTTCTT 1948
Db 1676 atgagagacagtgagggagagacagagactgaaattccggagctatttcagtttctt 1735
QY 1949 TTCCGTTTGTGCAATTTCACTTATGATACCGGCCAATGCTGTGTTGCTATTTTGAAC 2008
Db 1736 ttccgttttgtgcaatttcaacttatgataccggccaaatgctgtgttctattttggaac 1795
QY 2009 TCCCTTAGGGGATGCCCTCACTGCGCTTAAAGGGCCAGCCTGAGTGAGAGGAT 2068
Db 1796 tcccttaggggagtccttcaactgagcctataaaggccagcctgagctgagaggtat 1855
QY 2069 TCTGTGAGAGGATCAAGACAGACGCTGGACCTCGCACAGCCTCTCCACAGGTACC 2124
Db 1856 tctgtcagagga tcaagacagacgttggacctctgcagcctctccacaggtacc 1911

RESULT 4

AAQ46083
ID AAQ46083 standard; cDNA; 4299 BP.
XX
AC AAQ46083;
XX
DT 07-FEB-1994 (first entry)
XX
DE Sequence encoding a c-erbB-2 tumour antigen.
XX
KW Tumour antigen; c-erbB-2; glycoprotein; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..4299
FT /tag= a
XX
PN W09316185-A.
XX
PD 19-AUG-1993.
XX
PF 05-FEB-1993; 93WO-US01055.
XX
PR 06-FEB-1992; 92US-0831967.
XX
PA (CETU) CETUS ONCOLOGY CORP.
PA (CREA-) CREATIVE BIOMOLECULES INC.
XX
PI Houston LL, Huston JS, Oppermann H, Ring DB;
DR WPI; 1993-272889/34.
DR P-PSDB; AAR39568.
XX
PT New single chain Fv polypeptide binding to C-erbB-2 tumour antigen - for imaging or treating breast or ovarian cancer etc.
XX
PS Disclosure; pages 48-54; 87pp; English.
XX
CC c-erbB-2 refers to a protein antigen expressed on the surface of
CC tumour cells, such as breast and ovarian tumour cells, which is an
CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric
CC pt. of about 5.3 (see AAQ46083, AAR39568).
XX
SQ Sequence 4299 BP; 882 A; 1307 C; 1266 G; 844 T; 0 other;

Query Match 31.6%; Score 671.2; DB 14; Length 4299;
Best Local Similarity 95.4%; Pred. No. 1.7e-162;
Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
QY 228 ACGGACGACGAGATCACATGTCAGACCTGCCGCGGCTGGGGCATGTGCCACCA 287
Db 3075 accccagcaggggttctctgtccagacctgcccggcgctg999geatggtccacca 3134
QY 288 CAGGACCCGAGCTCATCTACAGAGGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 347
Db 3135 caggaccgcagctcatctaccagagtgcggtg99gacctgacactgagctgagacc 3194
QY 348 CTCTGAAGAGGAGGCGCCCGAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGT 407
Db 3195 ctctgaagagagggccccaggtctccactggcacctccgaagg99gctgctccgatgt 3254
QY 408 ATTTGATGTTGACTGGGATGGGCGAGCAAGGGGCTCAAGGCTCCCCACACATGA 467
Db 3255 atttgatggtgacctgggaatgggagccaaagg99gctgcaagcctcccccacatga 3314
QY 468 CCCAGCCCTCTACAGCGGTACAGTGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527
Db 3315 cccagccctctacagcgttacagtggagaccacacagtaacctcctctgagactga 3374
QY 528 TGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCCCTGAATATGTGAACAGCCAGATGT 587
Db 3375 tggctacgttggccccctgacctgagccccccagcctgaaatatgtgaaccagccagatgt 3434

QY 588 TCAGCCCGCCAGCCCTTCGCGCCCGAGAGGCGCCCTTCGCTGCTGCTGCCGACCTGCTGCTG 647
 Db 3435 tcggcccccagccccccttcgcccagagagggccctctgctgctgctgctgctgctgctgctgctg 3494
 QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGAGATGGGCTGCTCAAGACAGCTTTT 707
 Db 3495 cactctggaaagcccaagactctctcccagggagagatgggtcgtcaagagacgtttt 3554
 QY 708 TGCCCTTTGGGGTGGCGTGGAGAACCCGAG-ACTTGACACCCCGAGGAGGAGGAGCTGCCCC 766
 Db 3555 tgcctttgggggtgccttgagagaccccgagtagtacttgacaccccccagggagggctgcccc 3614
 QY 767 TCAGCCCGCCAGCTTCCTTCGCTTCAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 826
 Db 3615 tcagcccccagccctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 3674
 QY 827 GGACCCACAGAGCGGGGGGCTCCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 886
 Db 3675 ggaacccagagagcggggggctccacccagaccccttcaaaaggagacacctacggcagagaa 3734
 QY 887 CCAGAGTACCTGGGTCTGAGCTGCGAGCTGCGAGCTTAAGGCGCATATGCTGAGTGG 946
 Db 3735 ccagagtagctgggtctgagctgctgagctgagccagtggaaccagagagcccaagtcgcagaagcc 3794
 QY 947 ATGCCCTTGACCCAGG 962
 Db 3795 ctgatgtctctcagg 3810

RESULT 5
 AA231071
 ID AA231071 standard; DNA; 4473 BP.
 XX AA231071;
 DT 17-JAN-2000 (first entry)
 DE HER-2 nucleic acid sequence.
 XX HER-2; c-neu; ErbB2; transmembrane receptor; tyrosine kinase activity;
 KW epidermal growth factor receptor; EGFR; HER-1; cancer; breast cancer;
 KW ovarian cancer; gastric cancer; antisense oligonucleotide; expression;
 KW hyperproliferative disease; ss.
 XX Homo sapiens.
 OS
 PN WO9948906-A1.
 XX
 PD 30-SEP-1999.
 PF
 XX 25-MAR-1999; 99WO-US06492.
 PR 26-MAR-1998; 98US-0048804.
 XX (ISIS-) ISIS PHARM INC.
 PA (PENN-) PENN STATE RES FOUND.
 XX
 PI Bennett CF, Lipton A, Witters LM;
 XX WPI; 1999-610749/52.
 DR
 XX New antisense sequences used to treat hyperproliferative conditions,
 PT especially cancer -
 XX
 PS Examples; Page 38-39; 44pp; English.
 XX
 CC This is the human HER-2 polynucleotide sequence. The HER-2 gene also
 CC called c-neu and ErbB2, encodes a transmembrane receptor, with tyrosine
 CC kinase activity. HER-2 is related to the epidermal growth factor
 CC receptor (EGFR or HER-1). Aberrant HER-2 expression is present in a wide
 CC number of cancers, especially breast, ovarian and gastric cancers. This
 CC sequence is used in the invention to design 12-25 nucleotide

CC oligonucleotides that decrease the expression of human HER-2. The
 CC oligonucleotides of the invention (AA231067-231070) can also be used for
 CC modulating the expression of human epidermal growth factor receptor. The
 CC oligonucleotides are used to treat diseases or conditions associated with
 CC HER-2, particularly hyperproliferative diseases such as cancer.
 XX
 SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;

Query Match 31.6%; Score 671.2; DB 20; Length 4473;
 Best Local Similarity 95.4%; Pred. No. 1.8e-162;
 Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
 QY 228 AGCGCAGCAGAGAGATGTCACATGTCACAGCCCTGCCCGGCGCTGGGCGCTGGTCCACCA 287
 Db 3249 acccagcagggctctctctgctccagaccctgccccggcgctgggcatggtccacca 3308
 QY 288 CAGGACCCGAGCTCATCTACAGAGGTGGGGTGGGACCTTGACACTAGGGCTGGAGCC 347
 Db 3309 caggcaccgcagctcatctaccagagtggggtggacctgacactagggtgagcc 3368
 QY 348 CTCTGAAGAGAGGCCGCCAGGTCTCCACTGCGACCCCTCCGAAGGGGCTGCTCCGATGT 407
 Db 3369 ctctgaagagagggccccccaggtctccactggcaccctccgaagggtggtcgcagatgt 3428
 QY 408 ATTTGATGTTGACCTGGGAATGGGGCAGCCAAAGGGGCTGCAAAAGCTTCCCCACACATGA 467
 Db 3429 atttgatgtgacctgggaatggggcagccaaagggtgcaaaagctccccacacatga 3488
 QY 468 CCCCAGCCCTTACAGCGGTACAGTGAGGAGCCACACAGTACCCCTGCGCTCTGAGACTGA 527
 Db 3489 cccagccctctacagcgtacagtgagtgagagccacacagtacccctgcccctctgagactga 3548
 QY 528 TGGCTACGTTGCCCCCTGACCTGACCTGCGAGCCCGAGGCGCTGCAATATGTAACACGACGATGT 587
 Db 3549 tggctacgttgcctccctgacctgagccccagcctgataatgtgacccagccagatgt 3608
 QY 588 TCGGCCCCAGCCCTTCGCCCCGAGAGGGCCCTTCGCTGCTGCCCCGACCTGCTGGTGC 647
 Db 3609 tcggcccccagcccttcgcccagagagggccctctgctgctgctgctgctgctgctgctgctg 3668
 QY 648 CACTCTGAAGAGGCCAGACTCTCTCCCGAGGAGAGTGGGTGCTCAAGACGTTT 707
 Db 3669 cactctggaagggcccaagactctctcccaggggaagaatgggtcgtcaagagcgtttt 3728
 QY 708 TGCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACTTGACACCCCGAGGAGGAGGAGCTGCCCC 766
 Db 3729 tgcctttgggggtgccttgagaaaccccgagtagtacttgacaccccccagggagagtgcccc 3788
 QY 767 TCAGCCCGCCAGCTTCCTTCGCTTCAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 826
 Db 3789 tcagcccccagccctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 3848
 QY 827 GGACCCACAGAGCGGGGCTCCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 886
 Db 3849 ggaacccagagcggggggctccacccagacccctcaaaaggagacacctacggcagagaa 3908
 QY 887 CCCAGTACCTGGGTCTGAGCTGCGAGCTGCGAGCTTAAGGCGCATATGCTGAGTGG 946
 Db 3909 ccagagtagctgggtctgagctgctgagctgagccagtggaaccagagggcccaagtcgcagaagcc 3968
 QY 947 ATGCCCTTGACCCAGG 962
 Db 3969 ctgatgtgctcagg 3984

RESULT 6
 AAF24297
 ID AAF24297 standard; DNA; 9274 BP.
 XX
 AC AAF24297;
 XX
 DT 03-APR-2001 (first entry)

XX HER2 transgene plasmid construct.
 XX Human; HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
 KW antibody; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX WO200100244-A2.
 XX
 XX 04-JAN-2001.
 XX
 XX 23-JUN-2000; 2000WO-US17229.
 XX
 XX 25-JUN-1999; 99US-0141316.
 PR 16-MAR-2000; 2000US-0189844.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Erickson S, Schwall R;
 XX
 XX WPI; 2001-061962/07.
 DR P-PSDB; AAB60167.
 XX
 XX
 PT Treating tumors, particularly breast cancers, which overexpress an ErbB
 PT receptor and does not respond to an anti-ErbB antibody, comprises
 PT conjugating the antibody to a maytansinoid -
 XX
 XX Example 3; Fig 4; 92pp; English.
 XX
 XX The present invention provides a method of treating cancer by
 CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
 CC particular, the antibody is directed against ErbB2 (also known as HER2
 CC and p185neu). The method is particularly useful in the treatment of
 CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
 CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.
 XX
 XX Sequence 9274 BP; 2152 A; 2539 C; 2488 G; 2095 T; 0 other;

Query Match 31.5%; Score 670.4; DB 22; Length 9274;
 Best Local Similarity 93.8%; Pred. No. 3,7e-162;
 Matches 709; Conservative 0; Mismatches 46; Indels 1; Gaps 1;
 QY 228 ACGCAGCAGAGATCATCTCCAGACCCCTGCCCGGCGCTGGGGCATGTGCCACCA 287
 DB 4805 acccagcagggtcttctgtccagaccctgcccgggcgctgggcatgtgtccacca 4864
 QY 288 CAGCAGCGCAGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 347
 DB 4865 caggcaccgcagctcatctaccagagtg9gc9gt9gg9gacctgacactag9gct9ggagcc 4924
 QY 348 CTCTGAAGAGAGGCGCCCGAGGTCTCCACTGGCACCTCCCGAAGGGGTGGCTCCGATGT 407
 DB 4925 ctctgaagagagagcccccaggtctccactggcaccctccgaagg9gctgctccgatgt 4984
 QY 408 ATTTGATGGTGACCTGGGAATGGGGAGCCAGCAAGGGGTGCAAGCCCTCCCCACACATGA 467
 DB 4985 atttgatgtgacctgggaatgg9ggcagccaaagg9gctgcaaaagcctcccccacacatga 5044
 QY 468 CCCGAGCCTCTACAGCGGTACACTGAGGACCCACACAGTACCCCTGCCCTCTGAGACTGA 527
 DB 5045 cccagcctctacag9ggtacagt9ag9gacccacagctacccctgcccctgtgagactga 5104
 QY 528 TGGTACGTTGCCCCCTGACCTGCAGCGCCCGCAGCCTGAATATGTGAACACGAGCCAGATGT 587
 DB 5105 tggtagcttgcccccctgacctgagcccccagcctgaaatattgaaccagccagatgt 5164
 QY 588 TCGGCCCGAGCCCTTGGCCCGCAGAGGGGCCCTGTGCTGCTGCCCGCAGCTGTGTGTCG 647
 DB 5165 tcggcccccagcccttgcgcccagagaggccctctgctgctgctgctgctgctgctgctgctg 5224

QY 648 CACTCTGAAAGGCCCAAGACTCTCTCCCGAGGAAGATGGGTCTGTCAAAGAGGTTTT 707
 DB 5225 cactctgaaagggccaagactctctcccagggaagaatgggtcgtcaaaagcgtttt 5284
 QY 708 TGCCTTTGGGGTGGCTGGAGAGACCCCGAG-ACTTGACACCCCGAGGAGAGCTGCCCC 766
 DB 5285 tgcctttgggggtgcgtggagaaaccccgagtacttgacaccccgaggagctgcccc 5344
 QY 767 TCAGCCCCACACCTCTCTCTCTCAGCCCGAGCTTCGACACACCTCTATTACTGGGACCA 826
 DB 5345 tcagcccccacacctctctctcagcccgagccttcgacacacctctattctgggacca 5404
 QY 827 GGACCCACAGAGCGGGGGCTCCACCGACACCTTCAAGGAGGACACCTAGCGCAGAGAA 886
 DB 5405 ggaccaccagagcg9gg9gctccaccagcaccttcaaaaggagcacctagcgagagaa 5464
 QY 887 CCCAGAGTACCTGGTCTGGAGCTGCCAGTGTGAAGCCCTTAAGGCCCATATGCTGAGTGG 946
 DB 5465 cccagagtacctgggtctgcacgtgccagtgtaaccagaggcccaagtccgcagaaagcc 5524
 QY 947 ATGCTTTGACCCCGAGCGGGGATGGGGGAGACCTGT 982
 DB 5525 ctgatgtgtcctcagggagcagg9aaaggcgcctct 5560
 RESULT 7
 AAT01585
 ID AAT01585 standard; DNA; 4530 BP.
 XX
 XX AC AAT01585;
 XX
 XX 20-APR-1996 (first entry)
 XX
 XX Her-2/neu (ERBB2/c-erbB-2) gene sequence.
 DE
 XX Her-2/neu; Erb-B2; c-erbB-2; oncogene; DNA binding protein; HBBF;
 KW Erb-B2 promoter binding protein; tumour enhancer factor;
 KW breast cancer diagnosis; prognosis; antisense oligonucleotide;
 KW retro virus vector; gene therapy vector; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO9528485-A1.
 XX
 XX 26-OCT-1995.
 XX
 XX 19-APR-1995; 95WO-US04953.
 XX
 XX 19-APR-1994; 94US-0229515.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Raziuddin F, Sarkar FH;
 XX
 XX WPI; 1995-373800/48.
 XX
 XX New purified protein binding to the ERBB2 gene promoter - to induce
 PT cell proliferation, diagnostic of breast cancer, also related
 PT antibodies, nucleic acid, assays and methods for screening
 PT inhibitors.
 XX
 XX Disclosure; Page 52-54; 69pp; English.
 PS
 XX
 CC The Erb-B2 gene is one of the primary genes responsible for the
 CC transition of normal breast epithelial cells towards carcinoma in
 CC situ and the subsequent development of invasive and metastatic
 CC cancer. HBBF (see AAR77093-94), the Erb-B2 promoter binding protein,
 CC induces cell division on binding to the promoter. In a method for
 CC greater success in early identification and treatment of breast
 CC cancer, the initiation step for Erb-B2 gene activity is identified.
 CC This method involves determining the presence of HBBF in a biopsy
 CC from the subject, where the presence of HBBF (relative to its
 CC absence in a normal control) indicates the presence of cancer and

Db 3525 tggctacgtgtgccccctgaactgagccccagcgtgaatatgtgaaccagccagatgt 3584
Qy 588 TCGGGCCAGCCCTTGCCTCCGAGAGGGCCCTGCTGCTGCTGCCGACCTGCTGGTGC 647
Db 3585 tcggccccagcccccttcgccccagagggccctctgctgctgcccagacctgtggtgc 3644
Qy 648 CACTCTGAAAGGCCCAAGACTCTCTCCCGAGGGAAGTGGGTGCTCAAGAGCTTTT 707
Db 3645 cactctgaaaggccaaagactctctccccagggaagaatgggtcgcaaaagcgtttt 3704
Qy 708 TGCCTTTGGGGTGCCTGCGAGAACCCCGAG-ACCTGACACCCCGAGGAGCTGCCCC 766
Db 3705 tgcctttgggggtgcgtggagaaaccccgagtacttgacacccagggagagctgcgcc 3764
Qy 767 TCAGCCCCACCTCTCTCTGCTTCAGCCCGCTTCGACCAACCTCTATTACTGGGACCA 826
Db 3765 tcagccccacctctctgcttcagccccagcctctcgaacacctctattactgggacca 3824
Qy 827 GGACCCACAGAGCGGGGGTCCACCCAGCACCTTCAAAAGGACACCTAGCGCAGAAA 886
Db 3825 ggaccaccagagcgggggtccaccagacacttcaaaaggacacctagcgagagaa 3884
Qy 887 CCCAGAGTACCTGGGTCTGGAGCTGCCAGTGTGAAGCCTTAAGGGCCATATAGTGAGTGG 946
Db 3885 cccagagtacctgggtctggagctgcccagtggaaccagagggcccaagtcgcgagaaagcc 3944
Qy 947 ATGCTTGACCCACAGG 962
Db 3945 ctgatgtgtcctcagg 3960

RESULT 9

AAZ60815
ID AAZ60815 standard; DNA; 4530 BP.
AC
AC AAZ60815;
XX
XX 16-MAY-2000 (first entry)
XX
XX Nucleotide sequence of a cognate transgene of c-neu.
DE
XX
XX
KW Cognate transgene; CTG; tumourigenic; cellular immunogen; immunisation;
KW proto-oncogene; malignancy; allogenic cell; vaccine; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200004927-A1.
XX
XX PD 03-FEB-2000.
XX
XX PF 08-JUL-1999; 99WO-US15594.
XX
XX PR 24-JUL-1998; 98US-0093965.
XX
XX PA (UYAL-) UNIV ALLEGHNEY HEALTH SCI.
XX (HALP/) HALPERN M S.
XX (ENGL/) ENGLAND J M.
XX
XX PI Halpern MS, England JM;
XX
XX DR WPI; 2000-182543/16.
XX
XX PT Cellular immunogens comprising allogenic donor cells transfected with a
XX construct comprising a proto-oncogene cognate, useful as cancer
XX vaccines -
XX
XX PS Disclosure; Page 66-68; 77pp; English.
XX
XX CC The present sequence represents a cognate transgene (CTG) which is
XX rendered non-tumourigenic by deletion of amino acids.1-731. The CTG
XX is used in the course of the invention. The specification describes a
XX cellular immunogen for immunizing a host against the effects of the

CC product of a target proto-oncogene which is associated with a
CC malignancy. The cellular immunogen comprises allogenic cells
CC transfected with transgene construct comprising a transgene cognate
CC to target proto-oncogene and a strong promoter. The cellular immunogen
CC is useful for vaccinating a host against cancer by inserting the
CC transgene construct into the body of the host for the expression of
CC the transgene. The method of the invention is designed to target
CC mutation-driven non-self determinants. The cellular immunogens induce
CC reactivity for self-determinants in the over expressed product of tumour
CC associated and over expressed proto-oncogenes.
XX
SQ Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other;

Query Match 31.5%; Score 669.6; DB 21; Length 4530;
Best Local Similarity 95.2%; Pred. No: 4.5e-162;
Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
Qy 228 AGGGCAGCAGAAATCACATGTCCAGACCCCTGCCCGGGCGCTGGGGCATGGTCCACCA 287
Db 3225 accccagcagggctctctgtccagaccctgccccggcgctggggcatgtgtccacca 3284
Qy 288 CAGGCACCGCAGCTCATCTACAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCC 347
Db 3285 caggcaccgcagctcatctaccagagtgggcggtgggacctgacactagggctggagcc 3344
Qy 348 CTCTCAAGCAGGAGGCCGCCAGGCTCTCCACTGGCACCCCTCCGAAGGGGTGGCTCCGATGT 407
Db 3345 ctctgagagagggccccccaggtctccactggcacctccgaaagggtggctccgagtc 3404
Qy 408 ATTTGATGTGACCTGGGAATGGGGGCAGCCAAAGGGCTGCAAAAGCTCCCCACACATGA 467
Db 3405 attgatgtgacctggaaatggggcagcgaagggtgcaaaagcctccccacacatga 3464
Qy 468 CCCAGCCCTCTACAGCGGTACAGTGAGAGCCCAAGTACCCCTGCCCTCTGTGAGACTGA 527
Db 3465 cccagcctctacagcggtacagtgaggacccacagtacccctgcccctctgagactga 3524
Qy 528 TGGCTACGTGGCCCCCTGACCTGCAGCCCCCAGCGCTGAATATGTGAACCCAGCAGATGT 587
Db 3525 tggctacgtgtgccccctgaactgagccccagcgtgaatatgtgaaccagccagatgt 3584
Qy 588 TGGCCCCAGCCCCCTTGCCTCCGAGAGGGCCCTGCTGCTGCTGCCCGACCTGCTGGTGC 647
Db 3585 tcggccccagcccccttcgccccagagggccctctgctgctgcccagacctgtgtggtgc 3644
Qy 648 CACTCTGAAAGGCCCAAGACTCTCTCCCGAGGGAAGTGGGTGCTCAAGAGCTTTT 707
Db 3645 cactctgaaaggggccaaagactctctccccagggaagaatgggtcgcaaaagcgtttt 3704
Qy 708 TGCCTTTGGGGTGCCTGCGAGAACCCCGAG-ACCTGACACCCCGAGGAGCTGCCCC 766
Db 3705 tgcctttgggggtgcgtggagaaaccccgagtacttgacaccccgaggaggtgcgcc 3764
Qy 767 TCAGCCCCACCTCTCTCTGCTTCAGCCCCAGCCTTTCGACACCTCTATTACTGGGACCA 826
Db 3765 tcagccccacctctctctgcttcagccccagccttcgacaacctctattactgggacca 3824
Qy 827 GGACCCACAGAGCGGGGGTCCACCCAGCACCTTCAAAAGGACACCTACCGGCAGAGAA 886
Db 3825 ggaccaccagagcggggggtccacccagacacttcaaaaggacacctacggcagagaa 3884
Qy 887 CCCAGAGTACCTGGGTCTGGAGCTGCCAGTGTGAAGCCTTAAGGGCCATATAGTGAGTGG 946
Db 3885 cccagagtacctgggtctggagctgcccagtggaaccagagggcccaagtcgcgagaaagcc 3944
Qy 947 ATGCTTGACCCACAGG 962
Db 3945 ctgatgtgtcctcagg 3960

RESULT 10
AAAL4812


```

Db      3735 cccagagtacctgggtctgaggacgtgccagtgta 3768

RESULT 12
AA01912
ID   AAX01912 standard; DNA; 3768 BP.
AC     AC
AC     AC
XX    AAX01912;
DT      21-APR-1999 (first entry)
DE       Human HER-2/neu oncogene DNA.
XX
XX    HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW malignancy; treatment; tumour; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..3768
FT FT /*tag= a
FT FT /product= "HER-2/neu"
FT FT /note= "oncogene"
FT misc_feature 2026..3765
FT FT /*tag= b
FT FT /note= "region which elicits immune response"
XX
XX US5869445-A.
PN
XX
XX 09-FEB-1999.
PD
XX
XX 01-APR-1996; 96US-0625101.
XX
XX 01-APR-1996; 96US-0625101.
PR 17-MAR-1993; 93US-0033644.
PR 12-AUG-1993; 93US-0106112.
PR 31-MAR-1995; 95US-0411417.
XX
XX (UNIW ) UNIV WASHINGTON.
PA
XX Cheever MA, Disis ML;
PI
XX WPI; 1999-152835/13.
DR P-PADB; AA092406.
XX
XX Use of HER-2/new polypeptides - for eliciting an immune response to
PT an HER-2/new associated malignancy, particularly for treating or
PT preventing tumours
PT
XX Claim 1a; Column 23-32; 26pp; English.
XX
CC This sequence encodes the human HER-2/new oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immu
CC nese to HER-2/new protein. The polypeptide can stimulate T cells a
CC B cells to produce an immune response to the HER-2/new protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/new oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or reoccurrence.
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ

Query Match          31.4%; Score 667.6; DB 20; Length 3768;
Best Local Similarity 98.6%; Pred. No. 1.4e-161;
Matches 684; Conservative 0; Mismatches 9; Indels 1; Gaps

QY 228 ACGGCACGAGAACATCATGTCTCCAGACCCTGCCCGGGCGCTGGGGGCATGGTCCACCA 287
DB ||||| ||| |
DB 3075 acccagcagggttcttctgtccagaccctgcccggggcgtggggcatgtgccacca 3134

QY 288 CAGGCACCGCAGTCATCTACCCAGGAGTGCGCGGTGGGGACCTGACACTAGGGCTGGAGCC 347
DB ||||| ||| |
DB 3135 caggcacccgcagtcattaccagagatgacgactgggaacctgacactagggctggagccc 3194
```


Db 3735 cccagagtacctgggtctgtgacgtgccagtgta 3768

RESULT 14

AAA09455

ID AAA09455 standard; DNA; 3768 BP.

AC AAA09455;

XX 10-AUG-2000 (first entry)

XX Human heregulin 2 (Her2) coding sequence.

DE Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;

KW self-protein; cell-associated peptide antigen; foreign epitope;

KW cancer; breast cancer; prostate cancer; ss.

XX Homo sapiens.

OS WO200020027-A2.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-DK00525.

XX 05-OCT-1998; 98DK-0001261.

PR 20-OCT-1998; 98US-0105011.

XX (MEBI-) M & E BIOTECH AS.

PA Steina L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

XX Gautam A, Birk P, Karlsson G;

PI WPI: 2000-349917/30.

XX P-PSDB; AAY92620.

DR Inducing immune responses to weakly immunogenic, tumor associated

XX peptide antigens for the treatment of breast and prostate cancer

PT Claim 62; Page 187-193; 220pp; English.

PS The claims detail a method for inducing immune responses against weakly

XX immunogenic cell-associated peptide antigens (PA) such as those

CC associated with cancers (i.e. self-proteins), for example, human

CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or

CC fibroblast growth factor 8p (FGF8b). The method comprises effecting

CC simultaneous presentation by antigen producing cells (APCs) of the

CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)

CC group derived from the PA and/or at least 1 B-cell group derived from the

CC cell-associated PA; and (2) at least 1 first T helper cell group which is

CC foreign to the animal. Analogues of human PSM, human Her2 and

CC human/murine FGF8b comprising a substantial part of all known and

CC predicted CTL and B-cell epitopes of the respective PA and including at

CC least one foreign T helper epitope are also claimed. The method is used

CC to treat prostate, prostate/breast or breast cancer when the PA is human

CC PSM, FGF8b and Her2, respectively.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Query Match 31.3%; Score 666; DB 21; Length 3768;
Best Local Similarity 98.4%; Pred. No. 3.6e-161;
Matches 683; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

Qy 228 ACGGACGAGAGATACATGTCACAGACCTGCCCGCGCTGGGGCATGTGCCACCA 287

Db 3075 accccagcagggcttctgtccagaccctgcccggcgctggggcagtggtccacca 3134

Qy 288 CAGGACCGCAGCTCATCTACAGAGTGGGGGACCTGACACTAGGCTGGAGCC 347

Db 3135 caggcaccgagctcatctaccagagtggtgggacactgacactagggtggagcc 3194

Qy 348 CTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCTCCGAGGGGCTGGCTCCGATGT 407

Db 3195 ctctgaagagagggccccccaggtctccactggcacctccgaaggggctggctcagatgt 3254
Qy 408 ATTTCATGGTGACCTGGGAATGGGGGCGAGCCAGAGGGCTGCAAGCCCTCCCCACACATGA 467
Db 3255 atttgatggtgacctgggaatggggcagccaaggggtgcaaaagcctcccccacacatga 3314
Qy 468 CCCACAGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTGCCCTCTGAGACTGA 527
Db 3315 ccccaagcctctacagcgtacagtggagaccccaagtaacctgacctctgacctctgagactga 3374
Qy 528 TGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCCCTGAATGATGTAACCCAGCCAGATGT 587
Db 3375 tggctacgttgccccctgacctgacgtgagccccagcctgaatatgtgaaccagccagatgt 3434
Qy 588 TCGGCCCCAGCCCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCCCGCCGACCTGCTGGTGC 647
Db 3435 tgggccccagccccctctgccccggagggccccctgctgctgccccacactgctgggtgc 3494
Qy 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGGTCGTCAAAGAGCTTTT 707
Db 3495 cactctggaagggccaagactctctccccagggagaagtgggtcgtaaaagacgtttt 3554
Qy 708 TGCCCTTGGGGTGCCTGGGAGAACCCCGAG-ACCTTGACACCCCGAGGAGGAGCTGCCCC 766
Db 3555 tgcctttgggggtgcgtggagaaccccgagtacttgacaccccgaggagagctgcccc 3614
Qy 767 TCAGCCCCCACCCTCTCTCTGCTTTCAGCCCCAGCCCTTCGACAACTCTATTACTGGGACCA 826
Db 3615 tcagccccacactctctctgcttccagccagccttcgacacactctattactgggacca 3674
Qy 827 GGACCCACAGAGCGGGGGCTTCCACCCAGCACCTTCAAAGGGGACACCTTACCGCGAGAGAA 886
Db 3675 ggaccaccagagcggggggtccaccagcaaccttcaaaaggacacactacgagagaaa 3734
Qy 887 CCCAGAGTACCTGGGTCTGGGACGTGCCAGTGTGA 920
Db 3735 cccagagtacctgggtctgtgacgtgccagtgta 3768

RESULT 15

AAA89737

ID AAA89737 standard; cDNA; 3771 BP.

XX AAA89737;

XX 12-JAN-2001 (first entry)

XX Mouse Her-2/neu cDNA.

XX Mouse; Her-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

KW colon cancer; ss.

OS Mus sp.

XX WO200044899-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US02164.

XX 29-JAN-1999; 99US-0117976.

XX (CORI-) CORIXA CORP.

XX (SMIK) SMITHKLINE BEECHAM.

XX Cheever MA, Gheysen D;

XX WPI: 2000-505976/45.

DR P-PSDB; AAB21206.

XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:53:09 ; Search time 221.34 Seconds
(without alignments)
2174.328 Million cell updates/sec

Title: US-09-242-202A-22

Perfect score: 2125

Sequence: 1 GCCACCATGCCCTGACCT.....AGCCTCTCCACAGGTACCT 2125

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA.*
- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
 - 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
 - 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
 - 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
 - 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	671.2	31.6	4473	2	US-09-048-804-1
2	671.2	31.6	4473	4	US-09-036-105-26
3	669.6	31.5	4530	1	US-08-229-515A-9
4	669.6	31.5	4530	1	US-08-229-515A-9
5	667.6	31.4	3768	2	US-08-645-865-9
6	667.6	31.4	3768	2	US-08-625-101-1
7	495.8	23.3	3955	1	US-08-356-786-1
8	495.8	23.3	3955	1	US-08-229-515A-14
9	454.8	21.4	13254	1	US-08-645-865-14
10	454.8	21.4	13254	1	US-08-276-852-156
11	454.8	21.4	13254	1	US-08-276-852-170
12	454.8	21.4	13254	1	US-08-899-575-156
13	454.8	21.4	13254	1	US-08-899-575-170
14	454.8	21.4	13254	1	US-08-899-575-156
15	454.8	21.4	13254	5	US-08-899-575-170
16	454.8	21.4	13254	5	PCT-US95-08743-156
17	453.8	21.4	1905	1	US-08-594-469-9
18	453.8	21.4	1905	2	US-08-906-957-9
19	453.8	21.4	4410	1	US-08-594-469-1
20	453.8	21.4	4410	2	US-08-906-957-1
21	453.8	21.4	6151	5	PCT-US91-02954-12
22	453.8	21.4	8119	1	US-08-460-343B-1
23	453.8	21.4	8119	1	US-08-398-028B-1
24	453.8	21.4	8119	2	US-08-504-265B-1
25	453	21.3	2927	2	US-08-941-647A-1
26	453	21.3	2939	1	US-08-119-512-2
27	453	21.3	2939	1	US-08-488-015B-2

28 453 21.3 2939 3 US-08-814-412-17 Sequence 17, Appl
c 29 453 21.3 3003 6 5182260-18 Patent No. 5182260
c 30 453 21.3 3104 1 US-07-415-307A-1 Sequence 1, Appl
c 31 453 21.3 3104 1 US-08-371-320-1 Sequence 1, Appl
c 32 453 21.3 3130 4 US-09-038-141-1 Sequence 1, Appl
33 453 21.3 3249 1 US-08-507-455-4 Sequence 4, Appl
34 453 21.3 3301 2 US-08-447-430A-42 Sequence 42, Appl
35 453 21.3 3400 1 US-08-507-455-3 Sequence 3, Appl
36 453 21.3 3423 2 US-08-447-430A-40 Sequence 40, Appl
37 453 21.3 3474 2 US-08-447-430A-41 Sequence 41, Appl
38 453 21.3 3474 2 US-08-318-837-10 Sequence 10, Appl
39 453 21.3 3585 1 US-08-362-670B-9 Sequence 9, Appl
40 453 21.3 3585 3 US-08-333-576C-9 Sequence 9, Appl
41 453 21.3 3585 4 US-08-808-324-9 Sequence 9, Appl
42 453 21.3 3585 5 PCT-US94-14030A-9 Sequence 9, Appl
43 453 21.3 3623 2 US-07-989-847-13 Sequence 13, Appl
44 453 21.3 3623 4 US-08-469-411-13 Sequence 13, Appl
45 453 21.3 3632 1 US-07-745-382-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-048-804-1

; Sequence 1, Application US/09048804

; Patent No. 5968748

; GENERAL INFORMATION:

; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters

; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF

; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748RIS LLP

; STREET: One Liberty Place - 46th Floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: U.S.A.

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 Mb diskette

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/048,804

; FILING DATE: Herewith

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Paul K. Legaard

; REGISTRATION NUMBER: 38,534

; REFERENCE/DOCKET NUMBER: ISIS-2913

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4473 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Unknown

; ANTI-SENSE: No

US-09-048-804-1

Query Match 31.6%; Score 671.2; DB 2; Length 4473;

Best Local Similarity 95.4%; Pred. No. 7.4e-175;

Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 228 ACGGAGCAGACAGATCATGTCCAGACCCCTCCCGGGCGCTGGGGCATGGTCACCA 287

DB 3249 ACCCCAGCAGGGCTTCTTCTGTCCAGACCCCTCCCGGGCGCTGGGGCATGGTCACCA 3308

QY 288 CAGGACCCGACGCTCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGCTGAGCC 347
 Db 3309 CAGGACCCGACGCTCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGCTGAGCC 3368
 QY 348 CTCTGAAGAGGAGGCCCCCAGCTCTCCACTGCGACCTCCGAAGGGCTGGCTCCGATGT 407
 Db 3369 CTCTGAAGAGGAGGCCCCCAGCTCTCCACTGCGACCTCCGAAGGGCTGGCTCCGATGT 3428
 QY 408 ATTTGATGTGACCTTGGGAATGGGGAGCCAGCCAAAGGGCTGCAAGCCCTCCCCACACATGA 467
 Db 3429 ATTTGATGTGACCTTGGGAATGGGGAGCCAGCCAAAGGGCTGCAAGCCCTCCCCACACATGA 3488
 QY 468 CCCAGCCCTCTACAGCGGTACAGTGAAGACCCACAGTACCCCTGCGCTCTGAGACTGA 527
 Db 3489 CCCAGCCCTCTACAGCGGTACAGTGAAGACCCACAGTACCCCTGCGCTCTGAGACTGA 3548
 QY 528 TGGCTACGTTGCCCCCTGACCTGAGCCGCCAGCCCTGAATATGTGAACCCAGCAGATGT 587
 Db 3549 TGGCTACGTTGCCCCCTGACCTGAGCCGCCAGCCCTGAATATGTGAACCCAGCAGATGT 3608
 QY 588 TGGGCCCCAGCCCCCTTGGCCCCGAGAGGGCCCCCTGCTGCTGCTGCCACCTGCTGGTGC 647
 Db 3609 TGGGCCCCAGCCCCCTTGGCCCCGAGAGGGCCCCCTGCTGCTGCTGCCACCTGCTGGTGC 3668
 QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGGAGTGGGTCTCAAGAGCTTTT 707
 Db 3669 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGGAGTGGGTCTCAAGAGCTTTT 3728
 QY 708 TGCCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACVTTGACACCCAGGAGGAGCTGCCCC 766
 Db 3729 TGCCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTTCACCCAGGAGGAGCTGCCCC 3788
 QY 767 TCAGCCCCACCCCTCTCTGCTTCAGCCCCAGCCCTTCGACAACTCTATTACTGGGACCA 826
 Db 3789 TCAGCCCCACCCCTCTCTGCTTCAGCCCCAGCCCTTCGACAACTCTATTACTGGGACCA 3848
 QY 827 GGACCCACAGAGCGGGGGTCCACCCAGCACCCTTCAAGGGACACCTAGGCGAGAGAA 886
 Db 3849 GGACCCACAGAGCGGGGGTCCACCCAGCACCCTTCAAGGGACACCTAGGCGAGAGAA 3908
 QY 887 CCCAGTACCTGGGTCTGGAGCTGCCAGTGTGAAGCCCTTAAGGGCCATATGGTGAGTGG 946
 Db 3909 CCCAGTACCTGGGTCTGGAGCTGCCAGTGTGAAGCCCTTAAGGGCCATATGGTGAGTGG 3968
 QY 947 ATGCTTGACCCCGAGG 962
 Db 3969 CTGATGTGCTCTCAGG 3984

RESULT 2
 US-09-056-105-26
 ; Sequence 26, Application US/09056105
 ; Patent No. 6287569
 ; GENERAL INFORMATION:
 ; APPLICANT: KIPPS, THOMAS J.
 ; APPLICANT: WU, YUNQI
 ; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
 ; FILE REFERENCE: 233/221
 ; CURRENT APPLICATION NUMBER: US/09/056,105
 ; EARLIER FILING DATE: 1998-04-06
 ; EARLIER APPLICATION NUMBER: 60/043,467
 ; EARLIER FILING DATE: 1997-04-10
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 4473
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-056-105-26

Query Match 31.6%; Score 671.2; DB 4; Length 4473;

Best Local Similarity 95.4%; Pred. No. 7.4e-175;
 Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
 QY 228 ACGGCAGCAGAGAATACATGTCCAGACCCCTGCCCGGCGCTGGGCGATGTTCCACCA 287
 Db 3249 accccagcagggctctctctgtccagaccctgccccggcgctggggcattggtccacca 3308
 QY 288 CAGGACCCGACGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 347
 Db 3309 caggcaccgagctcatctaccagagtggggagtggggagctgacctagggctggagcc 3368
 QY 348 CTCTGAAGAGGAGGCCCCCAGCTCTCCACTGCGACCTCCGAAGGGCTGGCTCCGATGT 407
 Db 3369 ctctgaagagagggccccaggtctccactggcacctccgaaggggctggctccgatgt 3428
 QY 408 ATTTGATGTGACCTTGGGAATGGGGAGCCAAAGGGCTGCAAGCCCTCCCCACACATGA 467
 Db 3429 atttgaatggagctgggaatggggagcgaaggggctgcaagcctccccacacatga 3488
 QY 468 CCCAGCCCTCTACAGCGGTACAGTGAAGACCCACAGTACCCCTGCGCTCTGAGACTGA 527
 Db 3489 cccagccctctacaggggtacagtgaagacccacagtagccctgacctgagactga 3548
 QY 528 TGGCTACGTTGCCCCCTGACCTGAGCCGCCAGCCCTGAATATGTGAACCCAGCAGATGT 587
 Db 3549 tggctacgttgccccctgaactgcagccccagcctgaatatgtgaaccagcagatgt 3608
 QY 588 TCGGCCCCAGCCCCCTTGGCCCCGAGAGGGCCCCCTGCTGCTGCTGCCAGCCTGCTGGTGC 647
 Db 3609 tcggccccagcccccttcgccccgagagggccccctctgctgctgccccgacctgctggc 3668
 QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGGAGTGGGTCTCAAGAGCTTTT 707
 Db 3669 cactctggaaaggccccagactctctccccaggaagaatggggtctcaaaagacgtttt 3728
 QY 708 TGCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACVTTGACACCCAGGAGGAGCTGCCCC 766
 Db 3729 tgcctttgggggtggcgaggaaaccccgagtagtctgacacccagggagagctgcccc 3788
 QY 767 TCAGCCCCACCCCTCTCTGCTTCAGCCCCAGCCCTTCGACAACTCTATTACTGGGACCA 826
 Db 3789 tcagccccacccctctctctgcttcagccccagccttcgacacccctctattactgggacca 3848
 QY 827 GGACCCACAGAGCGGGGGTCCACCCAGCACCCTTCAAGGGACACCTAGGCGAGAGAA 886
 Db 3849 ggacccacagagcgggggctccacccagcaccctcaaaaggagacacctacggcagagaa 3908
 QY 887 CCCAGTACCTGGGTCTGGAGCTGCCAGTGTGAAGCCCTTAAGGGCCATATGGTGAGTGG 946
 Db 3909 cccagatgactgggtggtgagctgcccagtgcccagtggaaccagagggccaagtccgcagaagcc 3968
 QY 947 ATGCTTGACCCCGAGG 962
 Db 3969 ctgatgtgctctcagg 3984

RESULT 3
 US-08-229-515A-9
 ; Sequence 9, Application US/08229515A
 ; Patent No. 5518885
 ; GENERAL INFORMATION:
 ; APPLICANT: RAZIUDIN
 ; APPLICANT: SARKAR, FAZLUL H
 ; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
 ; TITLE OF INVENTION: NEOPLASTIC DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: usa
 ; ZIP: 30303

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/229,515A
 FILING DATE: 19 APR 1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: PERRYMAN, DAVID G
 REGISTRATION NUMBER: 33,438
 REFERENCE/DOCKET NUMBER: 1414.608
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 404-688-0770
 TELEFAX: 404-688-9880
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4530 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-229-515A-9

Query Match 31.5%; Score 669.6; DB 1; Length 4530;
 Best Local Similarity 95.2%; Pred. No. 2.1e-174;
 Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGATCATGTCACAGACCTGCCCGGGGCTGGGGGATGTCACACCA 287
 DB 3225 ACCCGACCGAGGCTTCTTGTGTGTCACAGACCTGCCCGGGGCTGGGGGATGTCACACCA 3284

QY 288 CAGGACCGGAGCTATCTACAGGAGTGCGGTGGGACCTTGACACTAGGGCTGGAGCC 347
 DB 3285 CAGGACCGGAGCTATCTACAGGAGTGCGGTGGGACCTTGACACTAGGGCTGGAGCC 3344

QY 348 CTCTGAGAGGAGGCCCCCAGGCTCTCACTGGGACCTCCGAAAGGGGCTGGCTCCGATGT 407
 DB 3345 CTCTGAGAGGAGGCCCCCAGGCTCTCACTGGGACCTCCGAAAGGGGCTGGCTCCGATGT 407

QY 408 ATTTGATGTGACCTGGGAATGGGCGAGCCAAAGGGCTGCAAAAGCTCCACACATGA 467
 DB 3405 ATTTGATGTGACCTGGGAATGGGCGAGCCAAAGGGCTGCAAAAGCTCCACACATGA 3464

QY 468 CCCCAGCCCTCTACAGCGGTACAGTGGAGGACCCACAGTACCCCTGCGCTCTGAGACTGA 527
 DB 3825 GGACCCACAGAGCGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGGCAGAGAA 3884

QY 887 CCAGAGTACCTGGTCTGGACGTGCCAGTGTGAAGCCTTAAGGGCCATATGTCGATGG 946
 DB 3885 CCAGAGTACCTGGTCTGGACGTGCCAGTGTGAAGCCTTAAGGGCCATATGTCGATGG 3944

QY 947 ATGCTTGTGACCCAGG 962
 DB 3945 CTGATGTGCTCAGG 3960

RESULT 4
 US-08-645-865-9
 ; Sequence 9, Application US/08645865
 ; Patent No. 5654406
 ; GENERAL INFORMATION:
 ; APPLICANT: RAZIUDIN
 ; APPLICANT: SARKAR, FAZLUL H
 ; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
 ; TITLE OF INVENTION: NEOPlastic DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: usa
 ; ZIP: 30303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/645,865
 ; FILING DATE: 14 MAY 1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERRYMAN, DAVID G
 ; REGISTRATION NUMBER: 33,438
 ; REFERENCE/DOCKET NUMBER: 1414.608
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 404-688-0770
 ; TELEFAX: 404-688-9880
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4530 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-645-865-9

Query Match 31.5%; Score 669.6; DB 1; Length 4530;
 Best Local Similarity 95.2%; Pred. No. 2.1e-174;
 Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGATCATGTCACAGACCTGCCCGGGGCTGGGGGATGTCACACCA 287
 DB 3225 ACCCGACCGAGGCTTCTTGTGTGTCACAGACCTGCCCGGGGCTGGGGGATGTCACACCA 3284

QY 288 CAGGACCGGAGCTATCTACAGGAGTGCGGTGGGACCTTGACACTAGGGCTGGAGCC 347
 DB 3285 CAGGACCGGAGCTATCTACAGGAGTGCGGTGGGACCTTGACACTAGGGCTGGAGCC 3344

QY 348 CTCTGAGAGGAGGCCCCCAGGCTCTCACTGGGACCTCCGAAAGGGGCTGGCTCCGATGT 407
 DB 3345 CTCTGAGAGGAGGCCCCCAGGCTCTCACTGGGACCTCCGAAAGGGGCTGGCTCCGATGT 407

QY 408 ATTTGATGTGACCTGGGAATGGGCGAGCCAAAGGGCTGCAAAAGCTCCACACATGA 467
 DB 3405 ATTTGATGTGACCTGGGAATGGGCGAGCCAAAGGGCTGCAAAAGCTCCACACATGA 3464

QY 468 CCCCAGCCCTCTACAGCGGTACAGTGGAGGACCCACAGTACCCCTGCGCTCTGAGACTGA 527
 DB 3825 GGACCCACAGAGCGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGGCAGAGAA 3884

NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3768 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..3768

OTHER INFORMATION: /note= "product = "cerB-b2""

Query Match 31.4%; Score 667.6; DB 2; Length 3768;
Best Local Similarity 98.6%; Pred. No. 6.7e-174;
Matches 684; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGATCATCTACAGAGTGGCGGTGGGACCTGCGGGCGTGGGGCATGGTCCACCA 287
DB 3075 ACCCCAGCAGGGCTTCTTGTGTCCAGACCTGCGGGCGGTGGGGCATGGTCCACCA 3134
QY 288 CAGGCACCGCAGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCC 347
DB 3135 CAGGCACCGCAGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCC 3194
QY 348 CTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGTGGCTCCGATGT 407
DB 3195 CTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGTGGCTCCGATGT 3254
QY 408 ATTTGATGGTGACCTGGGAATGGGGGCGAGCCAAAGGGGTGCAAGCCCTCCACACATGA 467
DB 3255 ATTTGATGGTGACCTGGGAATGGGGGCGAGCCAAAGGGGTGCAAGCCCTCCACACATGA 3314
QY 468 CCCACAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527
DB 3315 CCCACAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 3374
QY 528 TGGCTACGTTCGCCCTCCCTGACCTGACGAGCCCCCAGGCTGAATATGTGAACCCAGCAGATGT 587
DB 3375 TGGCTACGTTCGCCCTCCCTGACCTGACGAGCCCCCAGGCTGAATATGTGAACCCAGCAGATGT 3434
QY 588 TCGGCCCCAGCCCCCTTCGCCCCCCAGAGGGCCCTCTGCTGTGCTGCCCGACCTGTGGTGC 647
DB 3435 TCGGCCCCAGCCCCCTTCGCCCCCCAGAGGGCCCTCTGCTGTGCTGCCCGACCTGTGGTGC 3494

QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGTCGTCAAAGAGCTTTT 707
DB 3495 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGTCGTCAAAGAGCTTTT 3554
QY 708 TGCCCTTTGGGGGTGCCGTGGAGAACCCCGAG-ACITGACACCCCGAGGAGGAGCTGCCCC 766
DB 3555 TGCCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCC 3614
QY 767 TCAGCCGCCACCCCTCTCTGCGCTTCAGCCCGAGCTTCGACAACTCTATTACTGGGACCA 826
DB 3615 TCAGCCGCCACCCCTCTCTGCGCTTCAGCCCGAGCTTCGACAACTCTATTACTGGGACCA 3674
QY 827 GGACCCACCCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGAGCACCTTACGCGCAGAA 886
DB 3675 GGACCCACCCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGAGCACCTTACGCGCAGAA 3734
QY 887 CCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGA 920
DB 3735 CCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGA 3768

RESULT 7

US-08-229-515A-14

; Sequence 14, Application US/08229515A

; Patent No. 5518885

; GENERAL INFORMATION:

; APPLICANT: RAZIUDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN

; TITLE OF INVENTION: NEOPlastic DISEASE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEEDLE & ROSENBERG PC

; STREET: 127 Peachtree Street, Suite 1200

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: USA

; ZIP: 30303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/229,515A

; FILING DATE: 19 APR 1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: PERRYMAN, DAVID G

; REGISTRATION NUMBER: 33,438

; REFERENCE/DOCKET NUMBER: 1414.608

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 404-688-0770

; TELEFAX: 404-688-9880

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3955 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-229-515A-14

Query Match 23.3%; Score 495.8; DB 1; Length 3955;

Best Local Similarity 83.0%; Pred. No. 1.2e-126;

Matches 577; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 232 CAGCAGAGATCATGTCCAGACCTGCCCCGGCGTGGGGCATGTCCACACACAGG 291
DB 3110 CAGCAGGAGTATCTTCCCGCGACCTTACCACGAGCACTGGGAGCACACCCATAGAAGG 3169
QY 292 CACCAGCAGTCTATCTACACAGGAGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCTCT 351

Db 3170 CACCGAGCTGCTCCACAGGAGTGGTGGTACGTGACACTGGGCTCGAGCCCTCG 3229
QY 352 GAAGAGAGGCCCCAGGCTCCACTGGCACCCTCCGAAGGGCTGCTCCGATGATTT 411
Db 3230 GAAGAAGGGCCCCAGATCCACTGGCTCCCTCGAAGGGCTGCTCCGATGATTT 3289
QY 412 GATGGTACCTGGATGGGGGAGCGAGCGCTGCAAGGCTCCCCACACATGACCC 471
Db 3290 GATGGTACCTGGCAATGGGGTAAACAAAGGCTGCAGAGCTCTCTCCACATGACCTC 3349
QY 472 AGCCCTCTACAGCGGTACAGTGGAGGCCACACAGTACCCCTGCCCTCTGAGACTGATGC 531
Db 3350 AGCCCTCTACAGCGGTACAGTGGAGGCCACACATGATGCTGCTGCCCTCGAGCTGATGC 3409
QY 532 TAGCTGGCCCCCTGACCTGACGCCCCCAGCCCTGATATGTGAACAGCAGATGTTGG 591
Db 3410 TATGTTGCTCCCTGCTGAGCGCCAGCCCGAGTATGTGAACCAATCAGAGTTTCA 3469
QY 592 CCCAGCGCCCTTCGCCCCGAGAGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
Db 3470 CTTGAGCGCTCTTAACCCCGAGAGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 3529
QY 652 CTGGAAGGCCCTGCTCTCTCCAGGAGAAATGGGGTCTGCTGCTGCTGCTGCTGCTGCT 711
Db 3530 CTAGAAGAGCCCAAGACTCTCTCTCTGGAAGATGGGTTGTCAAGACGCTTTTGGC 3589
QY 712 TTTGGGGTGGCTGAGAACCCCG-AGACTTGACACCCCGAGGAGGCTGCTGCTGCTGCT 770
Db 3590 TTTGGGGTGGCTGAGAACCCCTGAATACTTAGTACGAGAGGAGGCTGCTGCTGCTGCT 3649
QY 771 CCCAGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
Db 3650 CCCAGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3709
QY 831 CCACAGAGCGGGGGCTCCACCCAGCAGCTTCAAGGAGACACCTTACGGCAGAGAACCCA 890
Db 3710 TCATCGGAGCGGGGGCTCCACCAAGTAACTTTGAAGGAGCCCGCCACTGCGAGAGACCT 3769
QY 891 GAGTACCTGGGTCTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
Db 3770 GAGTACCTAGGCTGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3804

RESULT 8

US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; City: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-14

Query Match 23.3%; Score 495.8; DB 1; Length 3955;
Best Local Similarity 83.0%; Pred. No. 1.2e-126;
Matches 57; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 232 CAGCAGAGATFACATGTCCAGACCCCTGCCCGGGCTCGGGGCGATGGTCCACACAGG 291
Db 3110 CAGCAGGAGTATCTTCTCCCGGACCTTACCCAGGCACTGGGAGCACAGCCCATAGNAGG 3169
QY 292 CACCGCAGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCT 351
Db 3170 CACCGCAGCTCGTCCACAGGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3229
QY 352 GAAGAGGAGGCCCGGCTCTCCACTGGCACCCTCCGAAAGGGCTGCTGCTGCTGCTGCTGCT 411
Db 3230 GAAGAAGGGCCCCCAGATCTCCACTGGCTCCCTCGGAAGGGCTGCTGCTGCTGCTGCTGCT 3289
QY 412 GATGGTGAACCTGGGAATGGGGGAGCCCAAGGGCTGCAAGCCCTCCACACATGATGATG 471
Db 3290 GATGGTGAACCTGGCAATGGGGTAAACAAAGGCTGCGAGAGCTCTCTCCACATGATGATG 3349
QY 472 AGCCCTCTACAGCGGTACAGTGGAGGCCACACAGTACCCCTGCTGCTGCTGCTGCTGCTGCT 531
Db 3350 AGCCCTCTACAGCGGTACAGTGGAGGCCACACATGATGATGATGATGATGATGATGATG 3409
QY 532 TAGCTGGCCCCCTGACCTGACGCCCCCAGCCTGATATGTGAACAGCAGATGTTGG 591
Db 3410 TATGTTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3469
QY 592 CCCAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
Db 3470 CTTGAGCGCTCTTAACCCCGAGAGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3529
QY 652 CTGGAAGGCCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCTGCTGCTGCTGCTGCTGCT 711
Db 3530 CTAGAAGAGCCCAAGACTCTCTCTCTGGAAGATGGGTTGTCAAGACGCTTTTGGC 3589
QY 712 TTTGGGGTGGCTGGAGAACCCCG-AGACTTGACACCCCGAGGAGGAGGCTGCCCTCTCAG 770
Db 3590 TTTGGGGTGGCTGGAGAACCCCTGAATACTTAGTACCGAGAGAGGAGGCTGCCCTCTCG 3649
QY 771 CCCAGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830
Db 3650 CCCAGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3709
QY 831 CCACAGAGCGGGGGCTCCACCCAGCAGCTTCAAGGAGACACCTTACGGCAGAGAACCCA 890
Db 3710 TCATCGGAGCGGGGGCTCCACCAAGTAACTTTGAAGGAGCCCGCCACTGCGAGAGACCT 3769
QY 891 GAGTACCTGGGTCTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
Db 3770 GAGTACCTAGGCTGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3804

RESULT 9

US-08-276-852-156/c
; Sequence 156, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A

```

; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-156

```

```

Query Match      21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
DB 6449 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 6390
QY 1471 CTGCTAATCCTGTTACCACTGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 1530
DB 6389 CTGCTAATCCTGTTACCACTGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6330
QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGGACGCTGCGGCTGAACGGGGGTTGCTGC 1590
DB 6329 GACTCAAGACGATAGTTACCGGATAAGGCGGACGCTGCGGCTGAACGGGGGTTGCTGC 6270
QY 1591 ACACAGCCCGCTGTTAGGCGAAGCACTACCGAACTGAGATACCTACAGCGTGAGCAT 1650
DB 6269 ACACAGCCCGCTGTTAGGCGAAGCACTACCGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 1651 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1710
DB 6209 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 6150
QY 1711 CTCGGNACAGGAGAGCGCAGGAGGAGCTTCCAGGGGAAAGCCCTGGTATCTTTATAGT 1770
DB 6149 GTCGGAACAGGAGAGCGCAGGAGGAGCTTCCAGGGGAAAGCCCTGGTATCTTTATAGT 6090
QY 1771 CCTGTGCGGGTTTCGCCACCTCTGACTTTGAGCGTTCGATTTTTTGTGATGCTCGTCAGGGGG 1830

```

```

DB 6089 CCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGG 6030
QY 1831 CGGAGCCTATGNAACGCCACCAACGCCAGCGCGCGGGG 1868
DB 6029 CGGAGCCTATGNAACGCCACCAACGCCAGCGCGCGGGG 5992

RESULT 10
; US-08-276-852-170
; Sequence 170, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-170

```

```

Query Match      21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
DB 6806 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 6865
QY 1471 CTGCTAATCCTGTTACCACTGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 1530
DB 6866 CTGCTAATCCTGTTACCACTGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6925
QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGGACGCGCTGCGGCTGAACGGGGGTTGCTGC 1590

```

```

Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTC 6985
QY 1591 ACACAGCCAGCTTGGAGCGAAGACCTTACACCGAAGTGGATACCTACACGCTGAGCAT 1650
Db 6986 ACACAGCCAGCTTGGAGCGAAGACCTTACACCGAAGTGGATACCTACACGCTGAGCAT 7045
QY 1651 TGAGAAAGCGGACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGAGG 1710
Db 7046 TGAGAAAGCGGACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGAGG 7105
QY 1711 GTCGGAACAGGAGCGCAGCAGGAGGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 1770
Db 7106 GTCGGAACAGGAGCGCAGCAGGAGGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 7165
QY 1771 CTTGTCGGGTTTCCGCACTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGGG 7225
Db 7166 CTTGTCGGGTTTCCGCACTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGGG 1830
QY 1831 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGGCGCGGGG 1868
Db 7226 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGGCGCGGGG 7263

```

RESULT 11

```

US-08-899-575-156/c
; Sequence 156, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 18-JUL-1994
; FILING DATE: 30-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-156
Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1411 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCAGCTACATACCTCGCT 1470
Db 6449 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCAGCTACATACCTCGCT 6390
QY 1471 CTGCTAATCTCTTACCAGTGGCTGCTGCCACTGCGATAAGTCTGTCTTACCGGGTTG 1530
Db 6389 CTGCTAATCTCTTACCAGTGGCTGCTGCCACTGCGATAAGTCTGTCTTACCGGGTTG 6330
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGCTGAACGGGGGTTTCGTGC 1590
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGCTGAACGGGGGTTTCGTGC 6270
QY 1591 ACACAGCCAGCTTGGAGCGAAGACCTTACCCGAACTGAGATACCTACACGCTGAGCAT 1650
Db 6269 ACACAGCCAGCTTGGAGCGAAGACCTTACCCGAACTGAGATACCTACACGCTGAGCAT 6210
QY 1651 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGCGAGGTATCCCGTAAGCGGAGG 1710
Db 6209 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGCGAGGTATCCCGTAAGCGGAGG 6150
QY 1711 GTCGGAACAGGAGCGCAGCAGGAGGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1770
Db 6149 GTCGGAACAGGAGCGCAGCAGGAGGCTTCCAGGGGAAACGCTGTATCTTTATAGT 6090
QY 1771 CTTGTCGGGTTTCCGCACTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 1830
Db 6089 CTTGTCGGGTTTCCGCACTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 6030
QY 1831 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGGCGCGGGG 1868
Db 6029 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGGCGCGGGG 5992

```

RESULT 12

```

US-08-899-575-170
; Sequence 170, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852

```


;
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACCTTCAGAACTCTGTAGCACCCTACATACCTCGCT 1470
Db 6806 GTGTAGCCGTAGTTAGGCCACCACCTTCAGAACTCTGTAGCACCCTACATACCTCGCT 6865
QY 1471 CTGCTAATCCTCTTACCAGTGTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 1530
Db 6866 CTGCTAATCCTCTTACCAGTGTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 6925
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCGTGGGTGAACGGGGGTTGTCG 1590
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCGTGGGTGAACGGGGGTTGTCG 6985
QY 1591 ACACAGCCCGCTTGGAGGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 6986 ACACAGCCCGCTTGGAGGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 7045
QY 1651 TGAGAAAGCCGACCTCTCCGAAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1710
Db 7046 TGAGAAAGCCGACCTCTCCGAAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 7105
QY 1711 GTCGGAACGAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGTGTATCTTTATAGT 1770
Db 7106 GTCGGAACGAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGTGTATCTTTATAGT 7165
QY 1771 CTTGTGCGGTTTTCGCCACCTCTGACTTGAGCGCTCGATTTTCTGTATGCTCGTCAGGGGG 1830
Db 7166 CTTGTGCGGTTTTCGCCACCTCTGACTTGAGCGCTCGATTTTCTGTATGCTCGTCAGGGGG 7225
QY 1831 CGGAGCCTATGGAAGAAACCCACGACGCGCGCGGGGG 1868
Db 7226 CGGAGCCTATGGAAGAAACCCACGACGCGCGCGGG 7263

RESULT 13
US-08-899-575-156/c
; Sequence 156, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of

;
; ADDRESSEE: Patent Counsel
; STREET: 10566 No. 5804440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 08/276,852
; FILING DATE: 30-SEP-1992
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-156

Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACCTTCAGAACTCTGTAGCACCCTACATACCTCGCT 1470
Db 6449 GTGTAGCCGTAGTTAGGCCACCACCTTCAGAACTCTGTAGCACCCTACATACCTCGCT 6390
QY 1471 CTGCTAATCCTCTTACCAGTGTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 1530
Db 6389 CTGCTAATCCTCTTACCAGTGTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 6330
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCGTGGGTGAACGGGGGTTGTCG 1590
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCGTGGGTGAACGGGGGTTGTCG 6270
QY 1591 ACACAGCCCGCTTGGAGGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 6269 ACACAGCCCGCTTGGAGGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 1651 TGAGAAAGCCGACCTCTCCGAAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1710
Db 6209 TGAGAAAGCCGACCTCTCCGAAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 6150
QY 1711 GTCGGAACGAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGTGTATCTTTATAGT 1770
Db 6149 GTCGGAACGAGGAGCGCACGAGGAGCTTCCAGGGGGAACCGCTGTGTATCTTTATAGT 6090
QY 1771 CTTGTGCGGTTTTCGCCACCTCTGACTTGAGCGCTCGATTTTCTGTATGCTCGTCAGGGGG 1830
Db 6089 CTTGTGCGGTTTTCGCCACCTCTGACTTGAGCGCTCGATTTTCTGTATGCTCGTCAGGGGG 6030

QY 1831 CGGAGCCTATGGAACACCGCAGCAACCGCGCGCGGGG 1868
 Db 6029 CGGAGCCTATGGAACACCGCAGCAACCGCGCGCGGAGG 5992

RESULT 14

US-08-899-575-170
 ; Sequence 170, Application US/08899575
 ; Patent No. 5804440
 ; GENERAL INFORMATION:
 ; APPLICANT: Burton, Dennis R
 ; APPLICANT: Barbas, Carlos F
 ; APPLICANT: Lerner, Richard A
 ; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
 ; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
 ; NUMBER OF SEQUENCES: 170
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: The Scripps Research Institute, Office of
 ; ADDRESSEE: Patent Counsel
 ; STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,
 ; STREET: Mail Drop TPC8
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/899,575
 ; FILING DATE: 24-JUL-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/276,852
 ; FILING DATE: 18-JUL-1994
 ; APPLICATION NUMBER: US 08/178,302
 ; FILING DATE: 30-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/954,148
 ; FILING DATE: 30-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fitting, Thomas
 ; REGISTRATION NUMBER: 34,163
 ; REFERENCE/DOCKET NUMBER: SCR1452P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-554-2937
 ; TELEFAX: 619-554-6312
 ; INFORMATION FOR SEQ ID NO: 170:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 13254 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-899-575-170

Query Match 21.4%; Score 454.8; DB 1; Length 13254;
 Best Local Similarity 99.6%; Pred. No. 4e-115;
 Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 1470
 Db 6806 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 6865
 QY 1471 CTGCTAATCTCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTGTCTTACCGGGTTG 1530
 Db 6866 CTGCTAATCTCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTGTCTTACCGGGTTG 6925
 QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGTTGCTGTC 1590

Db 6926 GACTCAAGACGATAGTTACCGGATAGGCGCAGCGGTTCGGGCTGAACGGGGGTTCTGTCG 6985
 QY 1591 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTAGATACCTACAGCGTGAACAT 1650
 Db 6986 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTAGATACCTACAGCGTGAACAT 7045
 QY 1651 TGAGAAAGCGCCACGCTTCCGAAAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAG 1710
 Db 7046 TGAGAAAGCGCCACGCTTCCGAAAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAG 7105
 QY 1711 GTGGAACAGAGAGCGCACGAGGAGCTTCCAGGGGGAAACGCTGTATCTTTATAGT 1770
 Db 7106 GTGGAACAGAGAGCGCACGAGGAGCTTCCAGGGGGAAACGCTGTATCTTTATAGT 7165
 QY 1771 CCTGTCCGGTTTCCGCACTCTGACTTGACGTCGATTTTGTGATGCTCGTCAGGGGGG 1830
 Db 7166 CCTGTCCGGTTTCCGCACTCTGACTTGACGTCGATTTTGTGATGCTCGTCAGGGGGG 7225
 QY 1831 CGGAGCCTATGGAACAAACGCGCAGCAACGCGCGCGGGG 1868
 Db 7226 CGGAGCCTATGGAACAAACGCGCAGCAACGCGCGCGGAGG 7263

RESULT 15

PCT-US95-08743-156/c
 ; Sequence 156, Application PC/TUS9508743
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
 ; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
 ; NUMBER OF SEQUENCES: 170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/08743
 ; FILING DATE: 11-JUL-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/276,852
 ; FILING DATE: 18-JUL-1994
 ; INFORMATION FOR SEQ ID NO: 156:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 13254 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; PCT-US95-08743-156

Query Match 21.4%; Score 454.8; DB 5; Length 13254;
 Best Local Similarity 99.6%; Pred. No. 4e-115;
 Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 1470
 Db 6449 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 6390
 QY 1471 CTGCTAATCTCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTGTCTTACCGGGTTG 1530
 Db 6389 CTGCTAATCTCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTGTCTTACCGGGTTG 6330
 QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGTTGCTGTC 1590
 Db 6329 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGTTGCTGTC 6270
 QY 1591 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTAGATACCTACAGCGTGAACAT 1650
 Db 6269 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTAGATACCTACAGCGTGAACAT 6210
 QY 1651 TGAGAAAGCGCCACGCTTCCGAAAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGG 1710

```
|||||
Db 6209 TGAGAAAGCGCCACGCTTCCGAAAGGGAGAAAAGGGGACAGGTATCCGGTAAAGCGGCAGG 6150
QY 1711 GTCCGGAACAGGAGAGCGGCACGAGGAGCTTCCAGGGGAAACGCCCTGTATCTTTATAGT 1770
Db 6149 GTCCGGAACAGGAGAGCGGCACGAGGAGCTTCCAGGGGAAACGCCCTGTATCTTTATAGT 6090
QY 1771 CCTGTCCGGGTTTCGGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1830
Db 6089 CCTGTCCGGGTTTCGGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 6030
QY 1831 CGGAGCCTATGGAACAAACGCCAGCAACGCGCGCGGGG 1868
Db 6029 CGGAGCCTATGGAACAAACGCCAGCAACGCGCGCGGAGG 5992
```

Search completed: January 17, 2002, 11:54:39
Job time: 17406 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:11 ; Search time 6501.33 Seconds
(without alignments)
3512.324 Million cell updates/sec

Title: US-09-242-202a-22
Perfect score: 2125
Sequence: 1 GCCACCATGGCCCTGACCT.....AGCCTCTCCACAGGTACCT 2125

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estpl:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	507.2	23.9	685	10 AU123871	AU123871
2	481.8	22.7	621	11 BG283493	BG283493 602407782
3	479.2	22.6	791	11 B1154872	B1154872 602902857
4	455.4	21.4	649	10 AA496412	AA496412 zv37a02.r
5	450.2	21.2	615	10 AV735664	AV735664 AV735664
6	449.8	21.2	527	10 AL043585	AL043585 DKF2p434G
7	449.8	21.2	571	10 AL044178	AL044178 DKF2p434P
8	449.8	21.2	579	10 AL043613	AL043613 DKF2p434H
9	449.8	21.2	616	10 AV735756	AV735756 AV735756
10	449.8	21.2	617	10 AV281661	AV281661 4A3A-P8G1
11	449.8	21.2	628	11 BF381364	BF381364 ASIR0004
12	449.8	21.2	954	10 AL044364	AL044364 DKF2p434C

c 13	448.8	21.1	629	10 AL593919	AL593919 AL593919
c 14	448.8	21.1	1004	10 AJ281480	AJ281480 4A3A-P4G8
c 15	448.4	21.1	609	10 AW701942	AW701942 uq93b02.y
c 16	448.2	21.1	703	10 AJ281437	AJ281437 4A3A-P4C3
c 17	448.2	21.1	800	10 AJ281449	AJ281449 4A3A-P4D5
c 18	447.8	21.1	568	10 AJ281376	AJ281376 4A3A-P2G2
c 19	447.8	21.1	579	10 AJ281320	AJ281320 4A3A-P4H1
c 20	447.8	21.1	700	10 AJ281616	AJ281616 4A3A-P8A1
c 21	446.6	21.0	1067	10 AU081137	AU081137 AU081137
c 22	445	20.9	808	10 AU176264	AU176264 AU176264
c 23	443.8	20.9	548	10 AJ281654	AJ281654 4A3A-P8F1
c 24	442.8	20.8	1070	10 AJ281552	AJ281552 4A3A-P6F1
c 25	441.8	20.8	526	10 AL043840	AL043840 DKF2p434A
c 26	441.2	20.8	498	10 AL039576	AL039576 DKF2p434D
c 27	438.8	20.6	615	10 AL044413	AL044413 DKF2p434E
c 28	438.4	20.6	741	11 BF299419	BF299419 2A4-6-11
c 29	437.8	20.6	734	10 AL039459	AL039459 DKF2p434O
c 30	437.2	20.6	1089	10 AU081124	AU081124 AU081124
c 31	435.8	20.5	480	13 C111G3	AJ226213 C1ona int
c 32	435.8	20.5	636	10 AJ281699	AJ281699 4A3A-P9E3
c 33	435.2	20.5	756	10 BE749097	BE749097 601123138
c 34	434.8	20.5	718	13 AG010489	AG010489 Homo sapi
c 35	433.6	20.4	959	10 BE749147	BE749147 601123315
c 36	431	20.3	759	10 BE749118	BE749118 6011233194
c 37	429.6	20.2	840	10 BE749178	BE749178 601123444
c 38	429.2	20.2	757	11 BF863156	BF863156 963041F03
c 39	427.2	20.1	786	11 B1155788	B1155788 602904360
c 40	425.8	20.0	462	11 BF834727	BF834727 CM3-HT096
c 41	425.8	20.0	630	10 AL042334	AL042334 DKF2p434O
c 42	425.2	20.0	784	13 AQ876119	AQ876119 V133E1 mt
c 43	423.8	19.9	590	10 AV613078	AV613078 AV613078
c 44	422.6	19.9	795	13 AQ876011	AQ876011 V132B5 mt
c 45	418.6	19.7	690	13 AG009464	AG009464 Homo sapi

ALIGNMENTS

RESULT 1
AU123871 AU123871 685 bp mRNA EST 23-OCT-2000
LOCUS AU123871 NT2RM2 Homo sapiens cDNA clone NT2RM2001211 5', mRNA
DEFINITION sequence.
ACCESSION AU123871
VERSION AU123871.1 GI:10948587
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES Location/Qualifiers
source
1..685
/organism="Homo sapiens"
/db_xref="taxon:9606"


```

VERSION      B1154872.1  GI:14614873
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM11090 row: b column: 15
              High quality sequence stop: 764.
FEATURES     Location/Qualifiers
              source
                1..791
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="IMAGE:5032238"
                /clone_lib="NIH_CGAP_Mam3"
                /tissue_type="tumor, gross tissue"
                /lab_host="DH10B"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert 2 kb. Library constructed by Life
                Technologies, catalog #12017-018. Investigators providing
                samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference
                for transgenic model: Xu et al., Nature Genetics 22, 37-43
                (1999). Note: this is a NCI_CGAP Library."
BASE COUNT   159 a 264 c 215 g 153 t
ORIGIN
Query Match      22.6%; Score 479.2; DB 11; Length 791;
Best Local Similarity 81.8%; Pred. No. 3.7e-109;
Matches 589; Conservative 0; Mismatches 128; Indels 3; Gaps 3;
QY 198 GGGGTGCTCTTGGGATCCTCATCAACGGCGGACAGAGATGTCACAGACCC 257
DB 11111111111111111111111111111111111111111111111111111
QY 68 GGAGTGTGTGATGCTGAAGAGTACTGGTACCCAGCAGGATGTTCTCCCCAGACCC 127
DB 11111111111111111111111111111111111111111111111111111
QY 258 TGCCCCGGCGCTGGGGCATGGTCCACACAGGACCGCAGCTCATCTACAGGAGTGG 317
DB 11111111111111111111111111111111111111111111111111111
QY 128 TGCCCTAGGTACTGGGAGCACAGGCCACCGCAGACACCCAGCTCGTCGGCCAGGATGG 187
QY 318 CGTGGGACCTGACACTGAGGCTGGAGCCCTCTGAAGAGAGGCCGCCAGGTCCTCACT 377
DB 11111111111111111111111111111111111111111111111111111
QY 188 CGGTGGTGTGAGTGTGACACTGGGCTGGAGCCCTCGGAAGAAGAGCCGCCAGATCTCCACT 247
DB 11111111111111111111111111111111111111111111111111111
QY 378 GGCACCTCCGAAGGGCTGGCTCGGATGTTTGTATGTGACCTGGGAATGGGGGACG 437
DB 11111111111111111111111111111111111111111111111111111
QY 248 GCTCCCTCCGAAGGGCTGGCTCGGATGTTTGTATGTGACCTGGGAGTGGGGTAAAC 307
DB 11111111111111111111111111111111111111111111111111111
QY 438 CAAGGGGTGCAAGACCTCCCCACACATGACCCACCGCCCTCTACAGCGGTACAGTGAGGA 497
DB 11111111111111111111111111111111111111111111111111111
QY 308 CAAAGGACTGCAGAGCCCTCTCCACATGACCTCAGACCTCTACAGCGGTACAGTGAGGA 367
DB 11111111111111111111111111111111111111111111111111111
QY 498 CCCACAGTACCCCTGCGCTTGAGACTGTAGGCTGACGTGGCCCTCCCTGACCTGACGCC 557
DB 11111111111111111111111111111111111111111111111111111
QY 368 TCCACATTAACCTCTGCCGCCGAGACTGATGGCTGACGTGCTGCCCTGGCCCTGACGCC 427
DB 11111111111111111111111111111111111111111111111111111
QY 558 CCAGCTCAATATGTGAACCCAGAGTGTCCGGCCCGCCAGCCCTTCGCCCCGAGAGGG 617
DB 11111111111111111111111111111111111111111111111111111
QY 428 CCAGCCCGAGTATGTGAACCCAGAGGTTGGGCTCAGTCTCCCTTGACCCCGAGAGGG 487
DB 11111111111111111111111111111111111111111111111111111
QY 618 CCCTCTGCTGTGTCGCCGACCTGTGTGGTGGCCACTCTGGAAAGGCCCAAGACTCTCTCCCC 677

```

```

Db 488 TCCTCCGCTCCCATCCGACCTGCTGGTACTACTAGAAAGACCAAGACTCTCTCTCC 547
QY 678 AGGGAAGAAATGGGTCGTCAAAGACGTTTTTGGCCCTTTGGGGTGGCGTGAGAACCCCG- 736
Db 548 TGGGAAAATGGGTTGTCAAAGACG-TTTTGGCCCTTTGGGGTGGCGTGAGAACCCCTGA 606
QY 737 AGACTTGACACCCAGGAGGAGCTGCCCTCAG-CCCCACACCTCTCTCTGCTTCAGCC 795
Db 607 ATACTAGACCCAGCAGCAGGACCTGCTCTCAGACCCCAACCTCTCTGCTTCAGCC 666
QY 796 CAGCCTTCGACCAACCTCTATTACTGGACACCCAGCAGCCAGCGGGGGTCCACCCA 855
Db 667 CAGCCTTTGACCAACCTCTATTACTGGACACCAACTCATCGCAGCAGGTCCTCCACCAA 726
QY 856 GCACCTTCAAAGGGACACCTAGCGGCAGAGAACCCAGACTACCTGGGTCTGGACGTGCCAG 915
Db 727 GTACCTTTGAAGGGAGCCCCCAATGACAGAGAACCTGACTAGGCTGGATGTGCCAG 786

RESULT 4
AA496412 649 bp mRNA EST 11-AUG-1997
LOCUS     zv37a02.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
DEFINITION IMAGE:755786 5' similar to gb:M11730 ERBB-2 RECEPTOR
            PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AA496412
VERSION    AA496412.1 GI:2229733
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (Bases 1 to 649)
AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B.,
            Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie
            ,T., Waterston,R. and Wilson,R.
            WashU-Merck EST Project 1997
            Unpublished (1997)
            Contact: Willson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 257.
FEATURES   Location/Qualifiers
            source
              1..649
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:755786"
              /clone_lib="Soares ovary tumor NbHOT"
              /sex="female"
              /tissue_type="ovarian tumor"
              /lab_host="DH10B (ampicillin resistant)"
              /note="Organ: ovary; Vector: pT73D (Pharmacia) with a
              modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
              strand cDNA was primed with a Not I - oligo(dT) primer [5'
              TGTTACCAATCTGAGTGGGAGCGCCGGTGTGTGTGTGTGTGTGTGTGT 3'],
              double-stranded cDNA was size selected, ligated to Eco RI
              adapters (Pharmacia), digested with Not I and cloned into
              the Not I and Eco RI sites of a modified pT73 vector
              (Pharmacia). Library constructed by Bento Soares and
              M.Fatima Bonaldo."
BASE COUNT 145 a 202 c 185 g 116 t 1 others
ORIGIN
Query Match      21.4%; Score 455.4; DB 10; Length 649;

```


Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.

This clone (DKF2p434G0127) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..527
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434G0127"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 113 a 158 c 137 g 119 t

Query Match 21.2%; Score 449.8; DB 10; Length 527;
Best Local Similarity 99.6%; Pred. No. 7.5e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
|||||
Db 474 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 415

QY 1471 CTGCTAATCTGTTACACGTGCTGCTCCAGTGGCGGATAGTCGTCTTACCGGGTTG 1530
|||||
Db 414 CTGCTAATCTGTTACACGTGCTGCTCCAGTGGCGGATAGTCGTCTTACCGGGTTG 355

QY 1531 GACTCAAGACGATAGTTACCGGATAGGCGCGCTCGGCTGGAACGGGGGTTCTGTCG 1590
|||||
Db 354 GACTCAAGACGATAGTTACCGGATAGGCGCGCTCGGCTGGAACGGGGGTTCTGTCG 295

QY 1591 ACACAGCCAGCTTGAGCGAGACGCTACACGGAAGTACGATACCTACAGCGTGAGCAT 1650
|||||
Db 294 ACACAGCCAGCTTGAGCGAGACGCTACACGGAAGTACGATACCTACAGCGTGAGCTA 235

QY 1651 TGAGAAAGCGCCACGCTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGG 1710
|||||
Db 234 TGAGAAAGCGCCACGCTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGG 175

QY 1711 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACGCCCTGTTATCTTTATAGT 1770
|||||
Db 174 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACGCCCTGTTATCTTTATAGT 115

QY 1771 CCTGTGCGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1830
|||||
Db 114 CCTGTGCGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 55

QY 1831 CGAGGCTATGGAAGAAACCCAGCAACGCGGCC 1863
|||||
Db 54 CGAGGCTATGGAAGAAACCCAGCAACGCGGCC 22

RESULT 7
AL04178/c
LOCUS AL04178 571 bp mRNA EST 29-FEB-2000
DEFINITION DKF2p434P0828_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL04178
VERSION AL04178.3 GI:5935954
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 571)
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Bloeker, et al.)
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5866789.
Contact: Bloeker H

MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.

This clone (DKF2p434P0828) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..571
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434P0828"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 122 a 168 c 149 g 132 t

Query Match 21.2%; Score 449.8; DB 10; Length 571;
Best Local Similarity 99.6%; Pred. No. 7.7e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
|||||
Db 463 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 404

QY 1471 CTGCTAATCTGTTACACGTGCTGCTGCCAGTGGCGGATAAGTCGTCTTACCGGGTTG 1530
|||||
Db 403 CTGCTAATCTGTTACACGTGCTGCTGCCAGTGGCGGATAAGTCGTCTTACCGGGTTG 344

QY 1531 GACTCAAGACGATAGTTACCGGATAGGCGCGCTCGGCTGGAACGGGGGTTCTGTCG 1590
|||||
Db 343 GACTCAAGACGATAGTTACCGGATAGGCGCGCTCGGCTGGAACGGGGGTTCTGTCG 284

QY 1591 ACACAGCCAGCTTGAGCGAGACGCTACACGGAAGTACGATACCTACAGCGTGAGCAT 1650
|||||
Db 283 ACACAGCCAGCTTGAGCGAGACGCTACACGGAAGTACGATACCTACAGCGTGAGCTA 224

QY 1651 TGAGAAAGCGCCACGCTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGG 1710
|||||
Db 223 TGAGAAAGCGCCACGCTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGG 164

QY 1711 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACGCCCTGTTATCTTTATAGT 1770
|||||
Db 163 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACGCCCTGTTATCTTTATAGT 104

QY 1771 CTGTGCGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1830
|||||
Db 103 CTGTGCGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 44

QY 1831 CGAGGCTATGGAAGAAACCCAGCAACGCGGCC 1863
|||||
Db 43 CGAGGCTATGGAAGAAACCCAGCAACGCGGCC 11

RESULT 8
AL043613/c

```
LOCUS      AL043613      579 bp      mRNA      EST      29-FEB-2000
DEFINITION DKFZp434H1527_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION  AL043613
VERSION     AL043613.1  GI:5423000
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 579)
AUTHORS     Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE       EST (Bloecker, et al.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferapitz 18a D-82152 Martinsried, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No r1 sequence available.
            This clone (DKFZp434H1527) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..579
Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="434 (synonym: htes3)"
    /tissue_type="testis"
    /dev_stage="adult"
    /lab_host="pH10B"
    /note="Vector: pSPort1; Site_1: NotI; Site_2: SalI"
BASE COUNT  128 a 171 c 152 g 128 t
ORIGIN
Query Match      21.2%; Score 449.8; DB 10; Length 579;
Best Local Similarity 99.6%; Pred. No. 7.7e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1411 GTGTAGCCCTAGTTAGGCCACCACTTCAGAACTCTGTAGCACCCTACATACCTCGCT 1470
DB 477 GTGTAGCCCTAGTTAGGCCACCACTTCAGAACTCTGTAGCACCCTACATACCTCGCT 418
QY 1471 CTGCTAATCTCTTACCAGTGGCTCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 1530
DB 417 CTGCTAATCTCTTACCAGTGGCTCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 358
QY 1531 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGCTGAACGGGGGTTGCGTCG 1590
DB 357 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGCTGAACGGGGGTTGCGTCG 298
QY 1591 ACACAGCCCCAGCTTGGAGCGAAGCACTTACACCGAAGTGTAGATACCTACACGCTGAGCAT 1650
DB 297 ACACAGCCCCAGCTTGGAGCGAAGCACTTACACCGAAGTGTAGATACCTACACGCTGAGCTA 238
QY 1651 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1710
DB 237 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 178
QY 1711 GTCGAAGCAGGAGCGCACGAGGAGCTTCCAGGGGAAACCGCTGTGTATCTTTATAGT 1770
DB 177 GTCGAAGCAGGAGCGCACGAGGAGCTTCCAGGGGAAACCGCTGTGTATCTTTATAGT 118
QY 1771 CCTGTCGGTTTCGCCACCTCTGACTTTCAGCGTCGATTTTGTGATCCTCGTCAGGGGGG 1830
DB 117 CCTGTCGGTTTCGCCACCTCTGACTTTCAGCGTCGATTTTGTGATCCTCGTCAGGGGGG 58
```

```
QY 1831 CGGAGCCTATGGAAAAACGCCAGCAACGCGGCC 1863
DB 57 CGGAGCCTATGGAAAAACGCCAGCAACGCGGCC 25

RESULT 9
AV735756/c
LOCUS      AV735756      616 bp      mRNA      EST      17-OCT-2000
DEFINITION AV735756 CB Homo sapiens cDNA clone CBMAGC03 5', mRNA sequence.
ACCESSION  AV735756
VERSION     AV735756.1  GI:10853337
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 616)
AUTHORS     Zhang,Q., Ye,M., Wu,X., Gu,J., Huang,Q., Zhou,J., Shen,Y., Han,Z.,
            Chen,S., Mao,M. and Chen,Z.
TITLE       Homo sapiens CB library cDNA clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zhu Chen
            Shanghai Institute of Hematology, Rui-Jin Hospital
            197 Rui-jin II Road, Shanghai 200025, P. R. China
            Tel: 86-21-64740490
            Fax: 86-21-64743206
            Email: mbs@shims.sh.cn
            This clone is available at Shanghai Hematology Institute in
            Shanghai.
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
FEATURES
source
1..616
Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="CBMAGC03"
    /clone_lib="CB"
    /tissue_type="cord blood"
    /cell_type="CD34+ hematopoietic stem/progenitor cell"
    /lab_host="BM25.8"
    /note="Vector: pBluescript; Site_1: EcoRI; The insert is
            cloned randomly with the EcoRI digestion"
BASE COUNT  125 a 182 c 165 g 143 t 1 others
ORIGIN
Query Match      21.2%; Score 449.8; DB 10; Length 616;
Best Local Similarity 99.6%; Pred. No. 7.8e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1411 GTGTAGCCCTAGTTAGGCCACCACTTCAGAACTCTGTAGCACCCTACATACCTCGCT 1470
DB 487 GTGTAGCCCTAGTTAGGCCACCACTTCAGAACTCTGTAGCACCCTACATACCTCGCT 428
QY 1471 CTGCTAATCTCTTACCAGTGGCTGTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 1530
DB 427 CTGCTAATCTCTTACCAGTGGCTGTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 368
QY 1531 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGCTGAACGGGGGTTGCTGTCG 1590
DB 367 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGCTGAACGGGGGTTGCTGTCG 308
QY 1591 ACACAGCCCCAGCTTGGAGCGAAGCACTTACACCGAAGTGTAGATACCTACACGCTGAGCAT 1650
DB 307 ACACAGCCCCAGCTTGGAGCGAAGCACTTACACCGAAGTGTAGATACCTACACGCTGAGCTA 248
QY 1651 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 1710
DB 247 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 188
QY 1711 GTCGGAAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACCGCTGTGTATCTTTATAGT 1770
```

Db 187 GTCCGAACAGGAGCGCACGAGGAGGCTTCACGGGGAAACGCTGTATCTTTATAGT 128
QY 1771 CTTGTCGGGTTTCGCCACCTCTGACTTCTGAGCTGCTGATTTTGTGATGCTCGTCAGSGGG 1830
Db 127 CTTGTCGGGTTTCGCCACCTCTGACTTCTGAGCTGCTGATTTTGTGATGCTCGTCAGSGGG 68
QY 1831 CGGAGCCTATGGAACAGCCAGCAACCGGCC 1863
Db 67 CGGAGCCTATGGAACAGCCAGCAACCGGCC 35

RESULT 10
AJ281661/c
LOCUS
DEFINITION
4A3A-P8G10-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P8G10, mRNA sequence.
ACCESSION
AJ281661
VERSION
AJ281661.1 GI:5929540
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE
1 (bases 1 to 617)
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.
and Kafatos, F.C.
TITLE
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
JOURNAL
MEDLINE
COMMENT
Location/Qualifiers
1. 617
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P8G10"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 136 a 175 c 162 g 144 t
ORIGIN

Query Match 21.2%; Score 449.8; DB 10; Length 617;
Best Local Similarity 99.6%; Pred. No. 7.8e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCGGTAGTTAGGCACACCTTCAGAACTCTGTAGCACCGCCCTACATACCTCGCT 1470
Db 488 GTGTAGCGGTAGTTAGGCACACCTTCAGAACTCTGTAGCACCGCCCTACATACCTCGCT 429
QY 1471 CTGCTAACTCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 1530
Db 428 CTGCTAACTCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 369
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCGGCTGGGGCTGAACGGGGGGTTTCGTGC 1590
Db 368 GACTCAAGACGATAGTTACCGGATAAGCGCGGCTGGGGCTGAACGGGGGGTTTCGTGC 309

QY 1591 ACACAGCCCGAGCTTGGAGCAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 308 ACACAGCCCGAGCTTGGAGCAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTA 249
QY 1651 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGG 1710
Db 248 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGG 189
QY 1711 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAGAAAGCGCTGTATCTTTATAGT 1770
Db 188 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAGAAAGCGCTGTATCTTTATAGT 129
QY 1771 CTTGTCGGGTTTCGCCACCTCTGACTTCTGAGCTGCTGATTTTGTGATGCTCGTCAGSGGG 1830
Db 128 CTTGTCGGGTTTCGCCACCTCTGACTTCTGAGCTGCTGATTTTGTGATGCTCGTCAGSGGG 69
QY 1831 CGGAGCCTATGGAACAGCCAGCAACCGGCC 1863
Db 68 CGGAGCCTATGGAACAGCCAGCAACCGGCC 36

RESULT 11
BF381364/c
LOCUS
DEFINITION
BF381364
stephensi cDNA 5', mRNA sequence.
ACCESSION
BF381364
VERSION
BF381364.1 GI:11370487
KEYWORDS
EST.
SOURCE
Anopheles stephensi.
ORGANISM
Anopheles stephensi.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE
1 (bases 1 to 628)
Xu, X., Qu, F., Song, G. and Xu, J.
AUTHORS
TITLE
The differentially expressing genes pool from Anopheles stephensi
related to infection with Plasmodium yoelii enriched by suppression
subtractive hybridization
Unpublished (2001)
JOURNAL
COMMENT
Contact: Xu Xiaochun; Qu Fengyi; Song Guan hong; Xu Jiannong
Department of Parasitology
Second Military Medical University
800 Xiangyin Rd., Shanghai, 200433, China
Tel: 86 021 25070276
Email: xcxu@mmu.edu.cn
Seq primer: T7
High quality sequence stop: 629
POLYA=No.

FEATURES
source
1. 628
/organism="Anopheles stephensi"
/strain="Hor"
/db_xref="taxon:30069"
/clone_lib="Mosquito Genes Pool Related Malaria Infection"
/sex="female"
/tissue_type="whole body"
/dev_stage="24 hours post-infection"

BASE COUNT 145 a 178 c 169 g 136 t
ORIGIN

Query Match 21.2%; Score 449.8; DB 11; Length 628;
Best Local Similarity 99.6%; Pred. No. 7.9e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCGGTAGTTAGGCACACCTTCAGAACTCTGTAGCACCGCCCTACATACCTCGCT 1470
Db 569 GTGTAGCGGTAGTTAGGCACACCTTCAGAACTCTGTAGCACCGCCCTACATACCTCGCT 510
QY 1471 CTGCTAACTCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 1530

```
Db 509 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTTG 450
QY 1531 GACTCAAGCAGTAGTTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGTTGCTGC 1590
Db 449 GACTCAAGCAGTAGTTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGTTGCTGC 390
QY 1591 ACACAGCCAGCTTTGAGGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 389 ACACAGCCAGCTTTGAGGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCTA 330
QY 1651 TGAGAAAGCGCCAGCTTCCGAAGCGAGAAAGCGGACAGTAGTATCCGTTAAGCGGCAGG 1710
Db 329 TGAGAAAGCGCCAGCTTCCGAAGCGAGAAAGCGGACAGTAGTATCCGTTAAGCGGCAGG 270
QY 1711 GTGGAAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTTGATCTTTATAGT 1770
Db 269 GTCGAAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTTGATCTTTATAGT 210
QY 1771 CCTGTGCGGTTTCCGACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTGTCAGGGGG 1830
Db 209 CCTGTGCGGTTTCCGACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTGTCAGGGGG 150
QY 1831 CGGAGCCTATGGAAGAAACCCAGCAACGCGGCC 1863
Db 149 CGGAGCCTATGGAAGAAACCCAGCAACGCGGCC 117
```

```
RESULT 12
AL044364/c
LOCUS DKF2p434C172.s1 434 bp mRNA EST 29-FEB-2000
DEFINITION DKF2p434C172 3', mRNA sequence.
ACCESSION AL044364
VERSION AL044364.1 GI:5432586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS Ansong,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Ansong, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansong W
MIPS Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKF2p434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: cloneerzpd.de.
FEATURES
source
1. .954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434C172"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN
```

Query Match 21.2%; Score 449.8; DB 10; Length 954;
Best Local Similarity 99.6%; Pred. No. 8.9e-102;

```
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1411 GTGTAGCGGTAGTTAGGCGACACCTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 1470
Db 851 GTGTAGCGGTAGTTAGGCGACACCTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 792
QY 1471 CTGCTAAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTTG 1530
Db 791 CTGCTAAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTTG 732
QY 1531 GACTCAAGCAGTAGTTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGTTGCTGC 1590
Db 731 GACTCAAGCAGTAGTTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGTTGCTGC 672
QY 1591 ACACAGCCAGCTTTGAGGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 671 ACACAGCCAGCTTTGAGGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCTA 612
QY 1651 TGAGAAAGCGCCAGCTTCCGAAGCGAGAAAGCGGACAGTAGTATCCGTTAAGCGGCAGG 1710
Db 611 TGAGAAAGCGCCAGCTTCCGAAGCGAGAAAGCGGACAGTAGTATCCGTTAAGCGGCAGG 552
QY 1711 GTCGAAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTTGATCTTTATAGT 1770
Db 551 GTCGAAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACCGCTTGATCTTTATAGT 492
QY 1771 CCTGTGCGGTTTCCGACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTGTCAGGGGG 1830
Db 491 CCTGTGCGGTTTCCGACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTGTCAGGGGG 432
QY 1831 CGGAGCCTATGGAAGAAACCCAGCAACGCGGCC 1863
Db 431 CGGAGCCTATGGAAGAAACCCAGCAACGCGGCC 399
```

```
RESULT 13
AL593919/c
LOCUS AL593919 629 bp mRNA EST 30-JUL-2001
DEFINITION AL593919 XGC-gastrula Silurana tropicalis cDNA clone Tgas003o10 5',
mRNA sequence.
ACCESSION AL593919
VERSION AL593919.1 GI:15005980
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 629)
AUTHORS Huckle,E., Taylor,R., McMurray,A., Ashurst,J.L., Zorn,A.M. and
Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxtion, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tgas003o10.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1. .629
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="Tgas003o10"
/dev_stage="gastrula (stages 10.5-13 mixed)"
/lab_host="Escherichia coli DH10B"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
```

FEATURES
source

into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

BASE COUNT 149 a 181 c 164 g 135 t
ORIGIN

Query Match 21.1%; Score 448.8; DB 10; Length 629;
Best Local Similarity 99.6%; Pred. No. 1.4e-101;
Matches 450; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCT 1470
Db 528 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCT 469
QY 1471 CTGCTAATCTGTTACCACTGCTGCTGCGAGTAAAGTCTGCTTACCGGGTTG 1530
Db 468 CTGCTAATCTGTTACCACTGCTGCTGCGAGTAAAGTCTGCTTACCGGGTTG 409
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGTGAACGGGGGTTGCTGC 1590
Db 408 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGTGAACGGGGGTTGCTGC 349
QY 1591 ACACAGCCAGCTTGGAGCGACGACCTACACCGAACTGAGATACCTACAGGTGAGCAT 1650
Db 348 ACACAGCCAGCTTGGAGCGACGACCTACACCGAACTGAGATACCTACAGGTGAGCTA 289
QY 1651 TGAGAAAGCGCCAGCTTCCGAAAGGAGAAAGCGGAGGAGTATCCGCTAAAGCGGCAGG 1710
Db 288 TGAGAAAGCGCCAGCTTCCGAAAGGAGAAAGCGGAGGAGTATCCGCTAAAGCGGCAGG 229
QY 1711 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCCCTGTTATCTTTATAGT 1770
Db 228 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCCCTGTTATCTTTATAGT 169
QY 1771 CTGTGCGGTTTGGCAGCTCTGACTTGAGCTGCGATTTTGTGATGCTGTCAGGGGG 1830
Db 168 CTGTGCGGTTTGGCAGCTCTGACTTGAGCTGCGATTTTGTGATGCTGTCAGGGGG 109
QY 1831 CGGAGCCTATGGAAGAACCCAGCAACGCGGC 1862
Db 108 CGGAGCCTATGGAAGAACCCAGCAACGCGGC 77

RESULT 14
AJ281480/c
LOCUS
DEFINITION
4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION
AJ281480
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
African malaria mosquito.
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
REFERENCE
1 (bases 1 to 1004)
AUTHORS
Dimopoulos, G., Casavant, J.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.
TITLE
Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE
20300950
COMMENT
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
LOCATION/Qualifiers
1. .1004
FEATURES
source
/organism="Anopheles gambiae"
/strain="4A r/r"

/db_xref="taxon:7165"
/clone_lib="4A3A-P4G8"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 252 a 262 c 244 g 244 t
ORIGIN

Query Match 21.1%; Score 448.8; DB 10; Length 1004;
Best Local Similarity 99.3%; Pred. No. 1.6e-101;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCT 1470
Db 475 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCCGCTACATACCTCGCT 416
QY 1471 CTGCTAATCTGTTACCACTGCTGCTGCGAGTAAAGTCTGCTTACCGGGTTG 1530
Db 415 CTGCTAATCTGTTACCACTGCTGCTGCGAGTAAAGTCTGCTTACCGGGTTG 356
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGTGAACGGGGGTTGCTGC 1590
Db 355 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGTGAACGGGGGTTGCTGC 296
QY 1591 ACACAGCCAGCTTGGAGCGACGACCTACACCGAACTGAGATACCTACAGGTGAGCAT 1650
Db 295 ACACAGCCAGCTTGGAGCGACGACCTACACCGAACTGAGATACCTACAGGTGAGCTA 236
QY 1651 TGAGAAAGCGCCAGCTTCCGAAAGGAGAAAGCGGAGGAGTATCCGCTAAAGCGGCAGG 1710
Db 235 TGAGAAAGCGCCAGCTTCCGAAAGGAGAAAGCGGAGGAGTATCCGCTAAAGCGGCAGG 176
QY 1711 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCCCTGTTATCTTTATAGT 1770
Db 175 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAAGCCCTGTTATCTTTATAGT 116
QY 1771 CTGTGCGGTTTGGCAGCTCTGACTTGAGCTGCGATTTTGTGATGCTGTCAGGGGG 1830
Db 115 CTGTGCGGTTTGGCAGCTCTGACTTGAGCTGCGATTTTGTGATGCTGTCAGGGGG 56
QY 1831 CGGAGCCTATGGAAGAACCCAGCAACGCGGC 1863
Db 55 CGGAGCCTATGGAAGAACCCAGCAACGCGGC 23

RESULT 15
AW701942
LOCUS
DEFINITION
uq93b02.y1 NCI CGAP Maml0 Mus musculus cDNA clone IMAGE:2538635 5' similar to gb:M11730 ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE
ACCESSION
AW701942
VERSION
AW701942.1 GI:7586086
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 609)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:03:53 ; Search time 599.86 Seconds
(without alignments)
17.151 Million cell updates/sec

Title: US-09-242-202a-22_copy_1_12
Perfect score: 12
Sequence: 1 GCCACCATGGCC 12

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	100.0	12	19	AAV21735
2	12	100.0	19	21	AAZ51587
3	12	100.0	19	21	AAZ51588
4	12	100.0	23	22	AAZ51589
5	12	100.0	23	22	AAZ51590
6	12	100.0	23	22	AAZ51591
7	12	100.0	25	14	AAQ41717
8	12	100.0	25	14	AAQ31109
9	12	100.0	27	14	AAQ52485
10	12	100.0	27	14	AAQ67276
11	12	100.0	27	16	AAQ101078
12	100.0	27	16	AAQ97620	

12	12	100.0	27	18	AAT84743	Primer BPI-23 for
13	12	100.0	27	19	AAV21728	PS3 cDNA RT-PCR pr
14	12	100.0	28	20	AAZ28010	Human lectomedin-2
15	12	100.0	28	21	AAZ96037	Human lectomedin-2
16	12	100.0	30	22	AAZ83005	Human MBSP2 cDNA c
17	12	100.0	31	20	AAV90571	Forward PCR primer
18	12	100.0	32	21	AAZ11968	Human truncated pl
19	12	100.0	32	22	AAZ77491	AlphaIb integrin
20	12	100.0	33	17	AAZ17090	Murine class II MH
21	12	100.0	33	19	AAZ62945	Murine signalling
22	12	100.0	33	20	AAZ02029	Murine signalling
23	12	100.0	33	21	AAZ60478	Murine signalling
24	12	100.0	33	21	AAA30367	Plasmid: TKHH2 PCR
25	12	100.0	33	21	AAZ14288	Murine MHC class I
26	12	100.0	33	21	AAZ86757	PCR primer #6048 f
27	12	100.0	33	21	AAZ38665	Ig-betaTh fusion g
28	12	100.0	33	22	AAZ27536	Primer #1. Uniden
29	12	100.0	33	22	AAZ83456	Primer. #1 used in
30	12	100.0	34	17	AAT09016	FLAP II-sense prim
31	12	100.0	35	21	AAZ63201	Adenovirus E2A cod
32	12	100.0	35	21	AAZ63004	Adenovirus E2A pri
33	12	100.0	35	21	AAZ52587	Wildtype early reg
34	12	100.0	35	22	AAZ80211	Adenovirus 5 E2A r
35	12	100.0	35	22	AAZ55483	PCR primer used to
36	12	100.0	35	22	AAZ58071	PCR primer used to
37	12	100.0	35	22	AAZ88882	Adenovirus E2A cod
38	12	100.0	36	19	AAZ17110	Adenovirus E2A pri
39	12	100.0	36	19	AAZ17111	Glutathione S-tran
40	12	100.0	36	22	AAZ56642	Glutathione S-tran
41	12	100.0	36	22	AAZ28415	Human alpha interf
42	12	100.0	37	22	AAZ23571	Human TGC cDNA PCR
43	12	100.0	38	19	AAZ32758	Asp2 coding sequen
44	12	100.0	38	22	AAZ75446	Human interferon-a
45	12	100.0	39	21	AAZ49595	Codon-optimised HP
			39	22	AAZ49595	Human delta-5-desa

ALIGNMENTS

RESULT 1
AAV21735
ID AAV21735 standard; cDNA; 12 BP.
XX
AC AAV21735;
XX
DT 17-AUG-1998 (first entry)
XX
DE Humanised vector 5' sequence acceptance site.
XX
KW Vector; vaccine; tumour; antigen; plasmid pITL; ss.
XX
OS Synthetic.
XX
PN WO9806863-A1.
XX
PD 19-FEB-1998.
XX
PF 14-AUG-1997; 97WO-US14306.
XX
PR 14-AUG-1996; 96US-0023931.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson EL, Nelson PJ;
XX
DR WPI; 1998-159552/14.
XX
PT Humanised polynucleotide vectors - comprising human derived promoter
PT and sequence acceptance site, used for the production of vaccines
PS Claim 13; Page 104; 125pp; English.
XX
CC This DNA sequence comprises a 5' acceptance site of novel humanised

CC polynucleotide vectors such as pITL (see AAV21724). Such vectors
 CC comprise a human-derived promoter or mammalian homologue and a
 CC sequence acceptance site that is designed to directionally accept
 CC sequence specific products from RT-PCR based cloning strategies
 CC via a unique site within an interrupted palindrome recognition
 CC sequence for a restriction endonuclease which is incorporated into
 CC the PCR primer. In this embodiment, the palindrome recognition
 CC sequence is for BglI. The 5' acceptance site reads on the positive
 CC strand, and includes a GCC sequence recognised by BglI and an ATG
 CC start codon. The 3' GCC moiety keeps the sequence in-frame and
 CC encodes Ser. The 3' acceptance site is given in AAV21736. The novel
 CC vectors are used to express target antigens, especially tumour
 CC antigens. They are non-replicating in mammalian cells but are
 CC capable of extended stable expression of target sequences, and
 CC generate immune responses to the target sequences with little or no
 CC immune response to the other vector components.

XX
 SQ Sequence 12 BP; 2 A; 6 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 12; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 1 gccaccatggcc 12
 |||||

RESULT 2

AAZ51587
 ID AAZ51587 standard; DNA; 19 BP.

XX AAZ51587;

XX 21-JUN-2000 (first entry)

DE PCR primer 5'ORF1 to amplify human Deltex (hZDX) coding region.

XX Notch signalling protein; Deltex; hZDX; human; cell differentiation;
 KW proliferation; ankyrin repeat; regulator; cytoskeletal; modulator; cancer;
 KW diagnosis; treatment; T lymphoblastic lymphoma; leukemia; lung carcinoma;
 KW adenocarcinoma; cervical neoplasia; PCR primer; ss.

XX Homo sapiens.

XX WO200012698-A1.

XX 09-MAR-2000.

XX 25-AUG-1999; 99WO-GB02802.

XX 31-AUG-1998; 98US-0098512.

XX (ZENE) ZENECA LTD.

XX Khoury-Christianson AM, Kays JS;

XX WPI; 2000-256639/22.

XX Novel human cytoplasmic Notch signaling protein and polynucleotide
 PT useful for identifying compounds that modulate Notch signaling protein
 PT and for diagnosis, treatment of leukemia, T lymphoblastic lymphoma -
 XX Example 2; Page 16; 60pp; English.

XX The present DNA sequence is the PCR primer, 5'ORF1 used to amplify the
 CC coding region from the start site of ORF1 through the common stop codon
 CC (479-2347), using human spleen cDNA library as the template. The 5' end
 CC PCR primer incorporates a Kozak consensus sequence prior to the ATG
 CC codon. The human cytoplasmic Notch signalling protein, Deltex (hZDX),
 CC is integral to cell differentiation and proliferation. It binds to the
 CC ankyrin repeat region of human Notch receptor and functions as a

CC regulator of Notch signalling pathway and has cytostatic activity. hZDX
 CC sequence is useful to identify compounds that modulate biological or
 CC pharmacological activity of Deltex Notch signalling protein. Specific
 CC antisense oligomers and dominant negative mutants are useful for
 CC therapeutic purposes. It is also used for diagnosis and treatment of
 CC conditions associated with aberrant Notch signalling, like cancers,
 CC including T lymphoblastic lymphoma, leukemia, adenocarcinoma, lung
 CC carcinoma and cervical neoplasia. The nucleotide sequence is useful
 CC for diagnostic assays to detect expression levels of hZDX.

XX Sequence 19 BP; 3 A; 9 C; 5 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 1 gccaccatggcc 12
 |||||

RESULT 3

AAZ51588

ID AAZ51588 standard; DNA; 19 BP.

XX AAZ51588;

XX 21-JUN-2000 (first entry)

DE PCR primer 5'ORF2 to amplify human Deltex (hZDX) coding region.

XX Notch signalling protein; Deltex; hZDX; human; cell differentiation;
 KW proliferation; ankyrin repeat; regulator; cytoskeletal; modulator; cancer;
 KW diagnosis; treatment; T lymphoblastic lymphoma; leukemia; lung carcinoma;
 KW adenocarcinoma; cervical neoplasia; PCR primer; ss.

XX Homo sapiens.

XX WO200012698-A1.

XX 09-MAR-2000.

XX 25-AUG-1999; 99WO-GB02802.

XX 31-AUG-1998; 98US-0098512.

XX (ZENE) ZENECA LTD.

XX Khoury-Christianson AM, Kays JS;

XX WPI; 2000-256639/22.

XX Novel human cytoplasmic Notch signaling protein and polynucleotide
 PT useful for identifying compounds that modulate Notch signaling protein
 PT and for diagnosis, treatment of leukemia, T lymphoblastic lymphoma -
 XX Disclosure; Page 16; 60pp; English.

XX The present DNA sequence is the PCR primer, 5'ORF2 used to amplify the
 CC coding region from the start site of ORF2 through the common stop codon
 CC (485-2347), using human spleen cDNA library as the template. The 5' end
 CC PCR primer incorporates a Kozak consensus sequence prior to the ATG
 CC codon. The human cytoplasmic Notch signalling protein, Deltex (hZDX),
 CC is integral to cell differentiation and proliferation. It binds to the
 CC ankyrin repeat region of human Notch receptor and functions as a
 CC regulator of Notch signalling pathway and has cytostatic activity. hZDX
 CC sequence is useful to identify compounds that modulate biological or
 CC pharmacological activity of Deltex Notch signalling protein. Specific
 CC antisense oligomers and dominant negative mutants are useful for
 CC therapeutic purposes. It is also used for diagnosis and treatment of
 CC conditions associated with aberrant Notch signalling, like cancers,
 CC including T lymphoblastic lymphoma, leukemia, adenocarcinoma, lung

CC carcinoma and cervical neoplasia. The nucleotide sequence is useful
CC for diagnostic assays to detect expression levels of hZDX.
XX
SQ Sequence 19 BP; 4 A; 10 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 12; DB 21; Length 19;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
| | | | | | | | | | | | | | |
Db 1 gccaccatggcc 12

RESULT 4

AAS01417
ID AAS01417 standard; DNA; 23 BP.

XX AAS01417;

AC AAS01417;

XX 04-JUL-2001 (first entry)

XX Mouse wild type alpha2delta-3 deletion mutant PCR primer #1.

XX Mouse; secreted calcium channel alpha2delta subunit; alpha2delta-2;
KW alpha2delta-3; alpha2delta-4; voltage-dependent calcium channel; VDCC;
KW gabapentin; scintillation proximity assay; SPA; nickel flashplate assay;
KW filter binding assay; wheat germ lectin flashplate assay; PCR primer; ss.

XX Mus musculus.

XX Key Location/Qualifiers

FT misc_feature 1..8

FT /tag= a

FT /note= "tagged region not present in template"

XX WO200119870-A2.

XX 22-MAR-2001.

XX 18-SEP-2000; 2000WO-EP09137.

XX 16-SEP-1999; 99US-0397550.

XX (WARN) WARNER LAMBERT CO.

XX Brown JP, Bertelli F;

XX WPI; 2001-235262/24.

XX Calcium channel alpha2delta subunits, useful in e.g. SPA assays,
PT Flashplate assays, Nickel Flashplate assays, Filter binding assays or
PT Wheat Germ Lectin Flashplate assays -

XX Example 12; Page 41; 160pp; English.

XX The present sequence for deletion mutant PCR primer #1 is used with
CC PCR primer #2 (AAS01418) to obtain mouse alpha2delta subunit deletion
CC mutant from the wild type mouse calcium channel alpha2delta-3 subunit
CC (AAU01026). The sequence is described in an invention relating to
CC truncated alpha2delta-2, alpha2delta-3 or alpha2delta-4 subunit soluble
CC proteins which retain their affinity for radioactively labelled
CC gabapentin. The alpha2delta subunit is 1 of the components of the
CC heteromultimeric voltage-dependent calcium channel (VDCC) complexes
CC present in neuronal and non-neuronal tissues including heart and
CC skeletal muscle. Numerous soluble forms of the human calcium channel
CC alpha2delta subunits (AAU01014-AAU01024 and AAU01032-AAU01038) and 5
CC soluble forms of the porcine calcium channel alpha2delta subunits
CC (AAU01027-AAU01031) are described. The secreted soluble alpha2delta
CC subunit may be used in assays e.g. scintillation proximity assay (SPA),
CC flashplate, nickel flashplate, filter binding or wheat germ lectin
CC flashplate assays to detect or measure the binding or interaction of a

CC ligand (e.g. gabapentin, L-Norleucine, L-Allo-Isoleucine, L-methionine,
CC L-Leucine, L-Isoleucine, L-Valine, Spermine and/or L-Phenylalanine) of
CC a calcium channel alpha2delta subunit.
XX
SQ Sequence 23 BP; 2 A; 10 C; 9 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
| | | | | | | | | | | | | | |
Db 3 gccaccatggcc 14

RESULT 5

AAF57554
ID AAF57554 standard; DNA; 23 BP.

XX AAF57554;

AC AAF57554;

XX 11-JUN-2001 (first entry)

XX Calcium channel alpha2delta subunit related primer.

XX Calcium channel alpha2delta subunit; alpha2delta-1; cerebral cortex;
KW nervous system disorder; pain; epilepsy; anxiety; PCR primer; ss.

XX Synthetic.

XX WO200120336-A2.

XX 22-MAR-2001.

XX 18-SEP-2000; 2000WO-EP09136.

XX 16-SEP-1999; 99US-0397549.

XX (WARN) WARNER LAMBERT CO.

XX Bertelli F, Brown JP, Dissanayake V, Suman-Chauhan N, Gee NS;

XX WPI; 2001-257902/26.

XX Competitive binding assay for screening ligands which bind a cerebral
PT cortical voltage-dependent calcium channel alpha2-delta-1 subunit,
PT where the ligands identified are useful for treating disorders of the
PT nervous system, including pain -

XX Disclosure; Page 104; 158pp; English.

XX The invention relates to a new method for screening ligands which bind a
CC cerebral cortical voltage-dependent calcium channel alpha2delta subunit,
CC preferably alpha2delta-1 subunit. The method comprises contacting a
CC secreted soluble recombinant alpha2delta-1 subunit with a ligand of
CC interest and a labelled compound which binds the subunit, followed by
CC measuring the level of binding of the labelled compound to alpha2delta-1
CC subunit. The method is useful for screening ligands, preferably
CC biologically active products that modulate a nervous system function,
CC which bind a cerebral cortical voltage-dependent calcium channel
CC alpha2delta-1 subunit. The ligands identified by the method are useful
CC for treating disorders of the nervous system, including pain, epilepsy
CC and anxiety. Sequences AAF57552-555 represent PCR primers related to
CC calcium channel alpha2delta subunits.

XX Sequence 23 BP; 2 A; 10 C; 9 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 22; Length 23;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||
 Db 3 gccaccatggcc 14

RESULT 6

AAQ41717
 ID AAQ41717 standard; DNA; 25 BP.
 XX AC
 XX AAQ41717;
 XX 25-AUG-1993 (first entry)
 XX Consensus eukaryotic ribosome binding site.
 XX Vaccine; cytotoxic T lymphocyte; CTL; influenza A virus;
 KW matrix protein; Ma; Pseudomonas exotoxin; cell recognition domain;
 KW translocation domain; anti-viral agent; ss.
 KW EP541335-A.
 PN
 XX 12-MAY-1993.
 XX 04-NOV-1992; 92EP-0310067.
 PF
 XX 08-NOV-1991; 91US-0792507.
 PR
 XX (MERI) MERCK & CO INC.
 PA
 XX Donnelly JJ, Friedman A, Hawe LA, Liu MA, Marshall MS;
 PI Montgomery DL, Oliff AA, Shi X, Ulmer J;
 XX WPI; 1993-154266/19.
 DR
 XX Recombinant DNA encoding bacterial toxin-antigen conjugates - are
 PT useful as vaccines against viral infections, tumours and
 PT parasites
 PT
 XX Example 6; Page 32; 8ipp; English.

PS The construct BS-PEMI encodes a hybrid protein comprising domains I
 CC and II from Pseudomonas exotoxin fused to amino acids 2-252 of the
 CC Influenza A virus matrix protein M1. The PE-derived portion of the
 CC hybrid protein allows internalisation of the protein by an antigen-
 CC presenting cell. The hybrid protein is then processed and an antigenic
 CC segment (i.e. the Influenza A virus matrix protein) is presented on
 CC the cell surface where it elicits an immune response. BSK-PEMI was
 CC made from BS-PEMI by replacing the 21bp XhoI/HindIII fragment with
 CC a fragment (i.e. AAQ41717) encoding a consensus eukaryotic ribosome
 CC binding site. The purpose of the construct was to increase the yields
 CC of in vitro translated PEMI protein.
 XX
 XX Sequence 25 BP; 6 A; 9 C; 8 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 14; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||
 Db 8 gccaccatggcc 19

RESULT 7

AAQ37109
 ID AAQ37109 standard; DNA; 25 BP.
 XX AC
 XX AAQ37109;
 XX 20-JUL-1993 (first entry)
 XX Consensus eukaryotic RBS.

XX PE; Pseudomonas exotoxin; influenza A virus; M1; matrix protein;
 KW BSK-PEMI; ribosome binding site; ss.
 XX Synthetic.
 XX EP532090-A.
 PN
 XX 17-MAR-1993.
 PD
 XX 02-SEP-1992; 92EP-0202660.
 PF
 XX 09-SEP-1991; 91US-0756249.
 PR
 XX (MERI) MERCK & CO INC.

PA Donnelly JJ, Friedman A, Hawe LA, Liu MA, Marshall MS;
 PI Montgomery DL, Oliff AI, Shi X, Ulmer J;
 XX WPI; 1993-087107/11.

XX Bacterial toxin-antigen protein conjugates - to elicit cytotoxic
 PT T-lymphocyte immune response, used for preventing viral
 PT infections, e.g. by influenza virus, HIV and human
 PT papilloma:virus
 XX
 PS Claim 24; Page 78; 85pp; English.
 XX
 CC Example 6 describes the construction of BSK-PEMI.
 CC BSK-PEMI was made from BS-PEMI by the replacement of the 21 bp
 CC XhoI/HindIII fragment with a 24 bp (sic) fragment encoding a
 CC consensus eukaryotic ribosome binding site. The purpose of the
 CC construct was to increase the yields of in vitro translated PEMI
 CC protein.
 XX
 XX Sequence 25 BP; 6 A; 9 C; 8 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 14; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||
 Db 8 gccaccatggcc 19

RESULT 8

AAQ52485
 ID AAQ52485 standard; DNA; 27 BP.
 XX AC
 XX AAQ52485;
 XX 03-JUN-1994 (first entry)
 XX Primer BPI-23.

XX Polymerase chain reaction; primer; amplify; PCR; plasmid; pING4533;
 KW pING4503; bactericidal/permeability-increasing protein; BPI; IgG;
 KW heavy; chain; pMB27; fusion protein; gram negative; renal failure;
 KW bacterial infection; endotoxin related shock; metabolic acidosis;
 KW disseminated intravascular coagulation; anemia; thrombocytopenia;
 KW leukopenia; adult respiratory distress syndrome; hypotension; fever;
 KW constant region; lipopolysaccharide complement binding;
 KW placental transfer; Fc receptor binding; ss.

XX Synthetic.

XX WO9323434-A.

XX 25-NOV-1993.

XX 19-MAY-1993; 93WO-US04754.

```
XX 19-MAY-1992; 92US-0885911.
XX (XOMA ) XOMA CORP.
XX Grinna LS, Horwitz A, Theofan G;
XX WPI; 1993-386485/48.
XX New fusion proteins for treating bacterial infections - comprising
XX a bactericidal-permeability-increasing protein and a immunoglobulin
XX heavy chain constant domain
XX Example 1; Page 20; 75pp; English.
XX The sequences given in AAQ52485-86 are primers which were used in the
XX construction of the rBPI-IgG fusion vector, pING4533. This plasmid
XX contains recombinant bactericidal/permeability-increasing protein
XX (rBPI)(1-191)ala132 with the initiating ATG in the context of the
XX consensus Kozak sequence given in AAQ52487. The rBPI fragment was fused
XX to a constant domain of an IgG heavy chain (HC). Plasmids such as
XX pING4533 may be used to produce fusion proteins which are useful for
XX the treatment of Gram negative bacterial infections and their sequelae
XX including endotoxin related shock and conditions associated with it,
XX such as disseminated intravascular coagulation, anemia,
XX thrombocytopenia, leukopenia, adult respiratory distress syndrome,
XX renal failure, hypotension, fever and metabolic acidosis. Proving
XX BPI or a fragment of it, as part of the fusion with an immuno-
XX globulin heavy chain constant region provides the potential advantages
XX of Fc receptor binding, bivalent binding to lipopolysaccharide
XX complement binding and increasing placental transfer.
XX Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
DB 10 gccaccatggcc 21

RESULT 9
AAQ67276
ID AAQ67276 standard; cDNA; 27 BP.
XX
AC AAQ67276;
XX
DT 29-MAR-1995 (first entry)
XX
DE rBPI(1-199)ala132 primer BPI-23.
XX
KW BPI; bactericidal-permeability-increasing protein; bactericide;
KW fusion protein; Gram-negative bacterium; infection; primer;
KW polymerase chain reaction; PCR; amplification; mutagenesis;
KW protein engineering; cysteine replacement analog; pING4533;
KW consensus Kozak translation sequence; rBPI(1-199)ala132; ss.
XX
OS Synthetic.
XX
PN WO9418323-A.
XX
PD 18-AUG-1994.
XX
PF 02-FEB-1994; 94WO-US01235.
XX
PR 02-FEB-1993; 93US-0013801.
XX
PA (XOMA ) XOMA CORP.
XX
PI Baltaian M, Burke D, Grinna L, Horwitz A, Theofan G;

XX WPI; 1994-279744/34.
XX Bactericidal-permeability-increasing protein analogues with
XX Cys132 or Cys135 replaced - also fusion proteins and C-terminally
XX truncated forms, have increased stability and are useful to treat
XX gram-negative bacterial infection
XX Disclosure; Page 11; 77pp; English.
XX Plasmid pING4533 contains a DNA insert encoding a C-terminally
XX truncated derivative of human mature BPI carrying a C132A mutation,
XX rBPI(1-199)ser135, in which the codon specifying the 5th amino
XX acid of the BPI signal peptide is placed in the context of the
XX consensus Kozak translation initiation sequence (AAQ67275) and the
XX first 4 amino acids of the BPI signal are removed. This is
XX accomplished by PCR amplification of BPI cDNA using primers BPI-23
XX (AAQ67276) and BPI-2 (AAQ67277).
XX Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
DB 10 gccaccatggcc 21

RESULT 10
AAQ01078
ID AAQ01078 standard; DNA; 27 BP.
XX
AC AAQ01078;
XX
DT 10-APR-1996 (first entry)
XX
DE rBPI(1-199)-Ala132 (with 27 amino acid signal) PCR primer BPI-23.
XX
KW bactericidal-permeability increasing protein; BPI; truncated; dimer;
KW endotoxic shock; heparin neutralisation; angiogenesis; inhibitor; ss.
XX
OS Synthetic.
XX
PN US5447913-A.
XX
PD 05-SEP-1995.
XX
PF 11-MAR-1994; 94US-0212132.
XX
PR 11-MAR-1994; 94US-0212132.
XX
PX (XOMA ) XOMA CORP.
XX
PY Ammons WS, Little RG;
XX
DR WPI; 1995-319904/41.
XX
OS Improved use of a bactericidal-permeability increasing protein -
XX for neutralisation of heparin and treatment of endo-toxaemia, by
XX admin. as covalently linked dimer of N-terminal fragment
XX Example 1; Column 15; 36pp; English.
XX Compositions of N-terminal fragments (i.e. amino acids 1-193 to
XX 1-199) of human bactericidal-permeability increasing (BPI) protein
XX are used to neutralise heparin and to treat endotoxaemia. Efficacy
XX of such compns. is improved when they comprise at least 50% of the
XX truncated rBPI in dimeric form. The present sequence is that of a
XX PCR primer used in the construction of a vector coding for rBPI(1-199)
XX in which native Cys at position 132 is replaced by Ala and the
```

CC first 4 amino acids of the BPI signal peptide are removed; the
 CC codon specifying the fifth amino acid of the signal sequence (Met
 CC at position -27) is placed in the context of a consensus Kozak
 CC translation initiation sequence. The truncated BPI is suitable
 CC for dimerisation in the presence of copper 2+ ions.
 XX
 SQ Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 16; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||
 Db 10 gccaccatggcc 21

RESULT 11
 AAQ97620
 ID AAQ97620 standard; cDNA; 27 BP.
 XX
 AC AAQ97620;
 XX
 DT 25-JAN-1996 (first entry)
 XX
 DE Probe/Primer for bactericidal/permeability increasing protein gene.
 XX
 KW Bactericidal permeability increasing protein; BPIP; infection;
 KW control; recombinant; ss.
 XX
 OS Synthetic.
 XX
 XX US5439807-A.
 PN
 XX 08-AUG-1995.
 PD
 XX 19-MAY-1992; 92US-0885501.
 XX
 PF 19-MAY-1993; 93US-0072063.
 PR
 XX 19-MAY-1992; 92US-0885501.
 PR
 PA (XOMA) XOMA CORP.
 XX
 XX Grinna LS;
 PI
 XX WPI; 1995-283094/37.
 XX
 DR Recovery of recombinant endotoxin binding protein - by culturing
 PT transformed cells in medium contg. cation exchange material, useful
 PT for treatment of bacterial infections.
 XX
 XX Disclosure; Column 21; 18pp; English.
 PS
 XX Recombinant endotoxin-binding protein comprising bactericidal/
 CC permeability increasing protein (BPIP) or its endotoxin N-terminal
 CC fragment, can be produced by growing genetically transformed host
 CC cells in a suitable medium so that BPIP is secreted into the medium.
 CC The BPIP is then isolated and can be used to control bacterial
 CC infections. The same method can be used where BPIP is replaced with
 CC a high density lipoprotein, limulus anti-LPS factor, tachyplesin or
 CC structurally related protein. This sequence is reproduced in the
 CC specification but is not referred to.
 XX
 SQ Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 16; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||
 Db 10 gccaccatggcc 21

RESULT 13
 AAV21728

Db 10 gccaccatggcc 21

RESULT 12
 AAT84743
 ID AAT84743 standard; DNA; 27 BP.
 XX
 AC AAT84743;
 XX
 DT 03-NOV-1997 (first entry)
 XX
 DE Primer BPI-23 for bactericidal permeability increasing protein DNA.
 XX
 KW Primer; polymerase chain reaction; PCR; amplification; Ig; BPI;
 KW immunoglobulin; heavy chain; constant region; plasmid pMB27;
 KW preparation; hybrid; fusion; bactericidal permeability increasing;
 KW treatment; Gram-negative; bacterium; bacteria; infection;
 KW endotoxic shock; disseminated intravascular coagulation; anaemia;
 KW anaemia; thrombocytopenia; thrombocytopenia; leukopenia;
 KW leukopenia; adult respiratory distress syndrome; ARDS;
 KW renal failure; hypotension; fever; metabolic acidosis; ss.
 XX
 OS Synthetic.
 XX
 XX US5643570-A.
 PN
 XX 01-JUL-1997.
 PD
 XX 19-MAY-1992; 92US-0885911.
 PF
 XX 19-MAY-1993; 93US-0064693.
 PR
 XX 19-MAY-1992; 92US-0885911.
 PR
 XX (XOMA) XOMA CORP.
 PA
 XX Grinna LS, Horwitz A, Theofan G;
 PI
 XX WPI; 1997-350186/32.
 DR
 XX Hybrid fusion protein for treating Gram-negative bacterial
 PT infections - comprising bactericidal/permeability increasing protein
 PT and immunoglobulin heavy chain constant domain
 PT
 XX Example 1; Columns 35-36; 31pp; English.
 PS
 XX The present sequence is primer for the PCR amplification of
 CC bactericidal permeability increasing (BPI) protein DNA from the
 CC plasmid pING4533. The DNA was used in the preparation of a hybrid
 CC fusion protein, comprising an amino-terminal BPI protein and a
 CC carboxy-terminal immunoglobulin heavy chain constant region. The
 CC fusion protein can be used to treat Gram-negative bacterial
 CC infections and their sequelae, e.g. endotoxic shock, and associated
 CC conditions, e.g. disseminated intravascular coagulation, anaemia,
 CC thrombocytopenia, leukopenia, adult respiratory distress syndrome
 CC (ARDS), renal failure, hypotension, fever or metabolic acidosis.
 CC The fusion protein provides the potential advantages of Fc receptor
 CC binding, bivalent lipopolysaccharide binding, complement binding
 CC and increased placental transfer.
 XX
 SQ Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 18; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||
 Db 10 gccaccatggcc 21

RESULT 13
 AAV21728

ID AAV21728 standard; cDNA; 27 BP.
 AC AAV21728;
 XX
 DT 17-AUG-1998 (first entry)
 XX
 DE P53 cDNA RT-PCR primer 1.
 XX
 KW Vector; vaccine; tumour; antigen; plasmid pITL-RHHR/neu;
 KW human; p53; cancer; RT-PCR; primer; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO9806863-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 14-AUG-1997; 97WO-US14306.
 XX
 PR 14-AUG-1996; 96US-0023931.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Nelson EL, Nelson PJ;
 XX
 DR WPI; 1998-159552/14.
 XX
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX
 PS Example 9; Page 47; 125pp; English.
 XX
 CC RT-PCR primer 1 and RT-PCR primer 2 (see AAV21729) were used to
 CC amplify a portion of p53 cDNA (see AAV21730). The primers
 CC include mutated sequences that allow the amplified product to
 CC be incorporated into vectors such as plasmid pITL (see AAV21724).
 CC Novel humanised vectors, which can be based on pITL, comprise a
 CC human-derived promoter (or mammalian homologue) which is functional
 CC in mammalian target tissue and cells and a sequence acceptance site
 CC which accepts cDNA products from RT-PCR cloning. The vectors are
 CC non-replicating in mammalian cells but are capable of extended
 CC stable expression of the target sequence, generating an immune
 CC response in immunised individuals. The vectors selectively elicit
 CC immune responses to the target sequences with little or no immune
 CC response to the other components of the vectors. The vectors are
 CC particularly useful in accommodating monomorphic and polymorphic
 CC nucleic acid sequences encoding tumor antigens via PCR technology.
 XX
 SQ Sequence 27 BP; 3 A; 13 C; 5 G; 6 T; 0 other;
 XX
 Query Match 100.0%; Score 12; DB 19; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCACCACATGGCC 12
 DB 5 gccaccatggcc 16
 XXXXXXXXXXXXXXX
 RESULT 14
 AA228010
 ID AA228010 standard; DNA; 28 BP.
 XX
 AC AA228010;
 XX
 DT 05-JAN-2000 (first entry)
 XX
 XX Human lectomedin-2 clone 2-1 amplifying primer JD#1.
 DE
 XX Human; 7-transmembrane receptor; lectomedin; lectin-binding; mucin;
 KW olfactomedin; cellular adhesion; atherosclerosis; gene therapy;
 XX

KW vascular disease; lectomedin-2; PCR primer; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9945111-A1.
 XX
 PD 10-SEP-1999.
 XX
 PF 04-MAR-1999; 99WO-US04676.
 XX
 PR 04-MAR-1998; 98US-0076782.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Hayflick JS;
 XX
 DR WPI; 1999-571596/48.
 XX
 PT New human lectomedin receptor polypeptide, used to identify specific
 PT binding partners for treating e.g. vascular disease
 XX
 PS Example 1; Page 36; 166pp; English.
 XX
 CC The invention provides purified and isolated human 7-transmembrane
 CC receptor lectomedin polypeptide or its fragments. The lectomedin
 CC polypeptide comprises extracellular lectin-binding, olfactomedin-like
 CC and mucin-like domains. The polypeptide can be produced by standard
 CC recombinant methodology. The polypeptide is involved in cellular adhesion
 CC and cytoplasmic metabolic pathways that are modulated by extracellular
 CC signaling. Specific binding to lectomedin-1 expressed on smooth muscle
 CC cells may be required for proliferation of these cells in
 CC atherosclerosis. The polypeptide is used to raise specific antibodies,
 CC and to identify specific binding agents that modulate (increase or
 CC decrease) its activity. The lectomedin nucleic acids are used as source
 CC of probes and primers, and of therapeutic antisense, ribozyme or triplex-
 CC forming agents, and in gene therapy to restore deficient lectomedin
 CC activity. Specific binding agents of lectomedin are used for treating
 CC diseases that involve lectomedin activity, e.g. vascular diseases such as
 CC atherosclerosis. The present sequence represents a primer for amplifying
 CC a human lectomedin-2 clone 2-1.
 XX
 SQ Sequence 28 BP; 6 A; 10 C; 6 G; 6 T; 0 other;
 XX
 Query Match 100.0%; Score 12; DB 20; Length 28;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCACCACATGGCC 12
 DB 14 gccaccatggcc 25
 XXXXXXXXXXXXXXX
 RESULT 15
 AAA96037
 ID AAA96037 standard; DNA; 28 BP.
 XX
 AC AAA96037;
 XX
 DT 26-JAN-2001 (first entry)
 XX
 XX Human lectomedin-2 cDNA 2.1 PCR primer JD#1.
 DE
 XX Human; lectomedin; seven transmembrane receptor protein;
 KW antiatherosclerotic; vasotropic; cytostatic; gene therapy;
 KW atherosclerosis; restenosis; vascular disease; peripheral neuropathy;
 KW cancer; nerve regeneration; renal cystic epithelium;
 KW uterine implantation; PCR primer; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX

```

PN WO200052039-A2.
XX
PD 08-SEP-2000.
XX
PF 03-MAR-2000; 2000WO-US05934.
XX
PR 04-MAR-1999; 99US-0262537.
XX
PA (ICOS-) ICOS CORP.
XX
PI Hayflick JS, Fox RD, Douangpanya J, Puri KD;
XX WPI; 2000-579274/54.
DR
XX
PT Isolated human seven transmembrane receptor lectomedin polypeptide or
PT its fragment, useful for identifying agents which can treat
PT atherosclerosis, restenosis or vascular disease -
XX
PS Example 1; Page 42; 241pp; English.
XX
CC The present sequence is a PCR primer which was used in a procedure for
CC isolating human seven transmembrane receptor lectomedin polypeptides.
CC The polypeptides have a characteristic extracellular structure including
CC lectin-binding, olfactomedin-like and mucin-like domains. Modulator
CC compounds that inhibit the binding of lectomedin to a binding partner
CC (galectin-3) may be used to treat conditions such as atherosclerosis,
CC restenosis, vascular disease, peripheral neuropathy, cancer, nerve
CC regeneration, and pathologies of the renal cystic epithelia and uterine
CC implantation. Nucleotide sequences encoding the lectomedin polypeptides
CC are useful in gene therapy. The lectomedin polynucleotides and
CC polypeptides may be used to identify lectomedin binding partner
CC compounds. The polynucleotides are useful for detecting lectomedin
CC expression in cells and for identifying genetic mutations in the
CC lectomedin genes. Antisense lectomedin nucleic acids are able to inhibit
CC expression of lectomedin genes.
XX
SQ Sequence 28 BP; 6 A; 10 C; 6 G; 6 T; 0 other;

Query Match 100.0%; Score 12; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
Db 14 gccaccatggcc 25

```

Search completed: January 17, 2002, 12:03:54
Job time: 17821 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:19 ; Search time 6501.33 Seconds

(without alignments)
19.834 Million cell updates/sec

Title: US-09-242-202a-22_COPY_1_12

Perfect score: 12

Sequence: 1 GCCACCATGGCC 12

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estfun:*

2: em_esthum:*

3: em_estin:*

4: em_estom:*

5: em_estpi:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_htc:*

10: gb_est1:*

11: gb_est2:*

12: gb_htc:*

13: gb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pin:*

18: em_gss_pro:*

19: em_gss_rod:*

20: em_gss_vrt:*

21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	100.0	49	11	BI078628
2	12	100.0	50	11	BF166238
3	12	100.0	56	11	BF166238
4	12	100.0	91	11	BG242301
5	12	100.0	99	10	BE259484
6	12	100.0	99	10	BE376022
7	12	100.0	101	11	BE400948
8	12	100.0	110	13	BE006903
9	12	100.0	115	11	AQ978348
10	12	100.0	117	11	BF852670
11	12	100.0	124	10	AA863181
12	12	100.0	124	11	BF092125

13	12	100.0	126	11	BG511069
14	12	100.0	127	10	AA214420
15	12	100.0	130	10	AA716853
16	12	100.0	131	11	N50245
17	12	100.0	133	11	BI152476
18	12	100.0	134	10	AU083189
19	12	100.0	135	10	AA784758
20	12	100.0	139	11	BG176372
21	12	100.0	141	11	BF533900
22	12	100.0	143	11	BG921952
23	12	100.0	143	11	BG982828
24	12	100.0	143	13	AZ225473
25	12	100.0	146	11	BG282693
26	12	100.0	148	10	BE256612
27	12	100.0	151	10	AA749848
28	12	100.0	151	10	BE366298
29	12	100.0	152	11	BF236756
30	12	100.0	157	10	AA750627
31	12	100.0	159	11	BF911674
32	12	100.0	160	10	AA279415
33	12	100.0	160	10	BE294646
34	12	100.0	161	10	BE270281
35	12	100.0	162	10	AI864485
36	12	100.0	163	10	BE266769
37	12	100.0	163	10	BE736479
38	12	100.0	164	10	AI715931
39	12	100.0	165	10	BE746616
40	12	100.0	165	13	AZ069587
41	12	100.0	166	10	BE269092
42	12	100.0	167	11	BF200800
43	12	100.0	168	11	BF752320
44	12	100.0	169	13	AZ086369
45	12	100.0	170	13	AZ314228

ALIGNMENTS

RESULT 1

BI078628/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI078628 49 bp mRNA EST 20-JUN-2001
602872729F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5004750 5',
mRNA sequence.
BI078628
BI078628.1 GI:14496958
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 49)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11043 row: i column: 07
High quality sequence stop: 49.
Location/Qualifiers
1. 49
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5004750"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"

FEATURES

source

```

/dev_stage="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      10 a      15 c      17 g      7 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 34 GCCACCATGGCC 23

RESULT 2
BF166238/c
LOCUS      50 bp      mRNA      EST      30-OCT-2000
DEFINITION 60177682F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4018606 5',
mRNA sequence.
ACCESSION  BF166238
VERSION     BF166238.1 GI:11046590
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 50)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9269 row: g column: 23
High quality sequence stop: 50.
FEATURES             Location/Qualifiers
     source          1..50
     /organism="Mus musculus"
     /strain="CZECH II (feral)"
     /db_xref="taxon:10090"
     /clone="IMAGE:4018606"
     /clone_lib="NCI_CGAP_Lu29"
     /tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
     /lab_host="DH10B"
     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      10 a      15 c      18 g      7 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 35 GCCACCATGGCC 24

/dev_stage="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      10 a      15 c      17 g      7 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 34 GCCACCATGGCC 23

RESULT 2
BF166238/c
LOCUS      50 bp      mRNA      EST      30-OCT-2000
DEFINITION 60177682F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4018606 5',
mRNA sequence.
ACCESSION  BF166238
VERSION     BF166238.1 GI:11046590
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 50)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9269 row: g column: 23
High quality sequence stop: 50.
FEATURES             Location/Qualifiers
     source          1..50
     /organism="Mus musculus"
     /strain="CZECH II (feral)"
     /db_xref="taxon:10090"
     /clone="IMAGE:4018606"
     /clone_lib="NCI_CGAP_Lu29"
     /tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
     /lab_host="DH10B"
     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      10 a      15 c      18 g      7 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 35 GCCACCATGGCC 24

```

```

RESULT 3
BG242301/c
LOCUS      56 bp      mRNA      EST      13-FEB-2001
DEFINITION 602354472F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4483048 5',
mRNA sequence.
ACCESSION  BG242301
VERSION     BG242301.1 GI:12752116
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 56)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10320 row: o column: 17
High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1..56
     /organism="Mus musculus"
     /strain="FVB/N"
     /db_xref="taxon:10090"
     /clone="IMAGE:4483048"
     /clone_lib="NCI_CGAP_Mam1"
     /tissue_type="tumor, biopsy sample"
     /dev_stage="3 months, virgin"
     /lab_host="DH10B"
     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      9 a      16 c      23 g      8 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 41 GCCACCATGGCC 30

RESULT 4
BG259484
LOCUS      91 bp      mRNA      EST      13-FEB-2001
DEFINITION 602378587F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4509501 5',
mRNA sequence.
ACCESSION  BG259484
VERSION     BG259484.1 GI:12769300
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 91)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: ATCC

```

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10389 row: m column: 22
 High quality sequence stop: 91.

FEATURES

Source
 1. 91
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4509501"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 16 a 30 c 39 g 6 t
 ORIGIN
 Query Match 100.0%; Score 12; DB 11; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||
 Db 30 GCCACCATGGCC 41

RESULT

BE376022 5
 LOCUS 99 bp mRNA EST 21-JUL-2000
 DEFINITION 601229459F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3593462 5',
 mRNA sequence.
 ACCESSION BE376022 GI:9321387
 VERSION BE376022.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 99)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM8766 row: e column: 15
 High quality sequence stop: 99.

FEATURES

Source
 1. 99
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3593462"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="3 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 21 a 27 c 44 g 7 t
 ORIGIN
 Query Match 100.0%; Score 12; DB 10; Length 99;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||
 Db 2 GCCACCATGGCC 13

RESULT

BE400948 99 bp mRNA EST 21-JUL-2000
 LOCUS AWB009.G08F000328 ITEC AWB Wheat Meiotic Stage Library Triticum
 DEFINITION aestivum cDNA clone AWB009.G08, mRNA sequence.

ACCESSION BE400948
 VERSION BE400948.1 GI:9360416
 KEYWORDS EST.
 SOURCE bread wheat.

ORGANISM

Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 99)
 AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier,
 S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
 Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
 Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
 Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M.,
 Sorrells, M., Warburton, M. and Wenzel, G.
 International Triticeae EST Cooperative (ITEC): Production of
 Expressed Sequence Tags for Species of the Triticeae
 Unpublished (2000)
 CONTACT: Langridge P
 Special Research Centre, Basic and Applied Plant Molecular Biology
 Waite Campus, University of Adelaide, Glen Osmond SA 5064 AUSTRALIA
 Tel: 61 8 8303 7368
 Fax: 61 8 8303 7102
 Email: plangridge@waite.adelaide.edu.au
 International Triticeae EST Cooperative (ITEC)
<http://wheat.pw.usda.gov/genome>.

JOURNAL

COMMENT

FEATURES

source
 1. 99
 /organism="Triticum aestivum"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="AWB009.G08"
 /clone_lib="ITEC AWB Wheat Meiotic Stage Library"
 /tissue_type="excised florets"
 /dev_stage="meiotic stage no later than metaphase I"
 /note="Vector: pSPORT 1 (Life Technologies cat. no.
 18248-013); Site 1: SalI; Library constructed in pSPORT 1.
 Directionally cloned using the Superscript Plasmid System
 for cDNA synthesis and plasmid cloning. M13 Reverse
 sequencing primer used to obtain 5' sequence data. 1.4
 kbp average insert size."

BASE COUNT 22 a 39 c 22 g 16 t
 ORIGIN

Query Match 100.0%; Score 12; DB 10; Length 99;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||
 Db 68 GCCACCATGGCC 79

```

RESULT 7
BG006903/c
LOCUS      101 bp      mRNA      EST      24-JAN-2001
DEFINITION IL5-GN0239-271100-273-g10_1 GN0239 Homo sapiens cDNA, mRNA
sequence.
ACCESSION  BG006903
VERSION     BG006903.1  GI:12450544
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 101)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0239-
            271100-273-g10_1&t3=2000-11-27&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 101.
FEATURES   Location/Qualifiers
            source          1..101
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_lib="GN0239"
                        /dev_stage="Adult"
                        /note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
                        ; Site_2: SmaI; A mini-library was made by cloning
                        products derived from ORESTES PCR (U.S. Letters Patent
                        application No. 196,716 - Ludwig Institute for Cancer
                        Research) profiles into the puc 18 vector. Reverse
                        transcription of tissue mRNA and cDNA amplification were
                        performed under low stringency conditions."
BASE COUNT 30 a      28 c      24 g      19 t
ORIGIN
Query Match      100.0%; Score 12; DB 11; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 81 GCCACCATGGCC 70

RESULT 8
AQ978348
LOCUS      110 bp      DNA      GSS      29-JAN-2000
DEFINITION RPCI-23-321E6.TJ RPCI-23 Mus musculus genomic clone RPCI-23-321E6,
            DNA sequence.
ACCESSION  AQ978348
VERSION     AQ978348.1  GI:6810649
KEYWORDS   GSS.
SOURCE     house mouse.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 115)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.

```

```

ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 110)
AUTHORS    Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
TITLE      Mouse BAC End Sequences from Library RPCI-23
JOURNAL    Unpublished (1999)
COMMENT    Other_GSSs: RPCI-23-321E6.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 321 row: E column: 6
            Seq primer: SP6
            Class: BAC ends.
FEATURES   Location/Qualifiers
            source          1..110
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="RPCI-23-321E6"
                        /clone_lib="RPCI-23"
                        /sex="Female"
                        /lab_host="DH10B"
                        /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                        EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                        brain genomic DNA was isolated and partially digested
                        with a combination of EcoRI and EcoRI Methylase. Size
                        selected DNA was cloned into the pBACe3.6 vector at the
                        EcoRI sites. The ligation products were transformed into
                        DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 29 a      22 c      28 g      31 t
ORIGIN
Query Match      100.0%; Score 12; DB 13; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 8 GCCACCATGGCC 19

RESULT 9
BG980184/c
LOCUS      115 bp      mRNA      EST      12-JUN-2001
DEFINITION PK3-CN0100-230101-002-h05 CN0100 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG980184
VERSION     BG980184.1  GI:14382919
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 115)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.

```

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM3&t2=PM3-CN0100-230101-002-h05&t3=2001-01-23&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 114.

FEATURES
source
1. .115
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0100"
/dev_stage="Adult"
/note="Organ: colon_normal; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 34 a 23 c 32 g 26 t
ORIGIN
Query Match 100.0%; Score 12; DB 11; Length 115;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 105 GCCACCATGGCC 94

RESULT 10

LOCUS BF852670 117 bp mRNA EST 16-JAN-2001

DEFINITION CM3-EN0194-151200-563-g04 EN0194 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF852670

VERSION BF852670.1 GI:12240414

KEYWORDS EST.

SOURCE human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 117)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

REFERENCE
AUTHORS Simpson A.J.G.
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001.

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-EN0194-151200-563-g04&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 117.

FEATURES
source
1. .117
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0194"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 18 a 36 c 36 g 27 t
ORIGIN
Query Match 100.0%; Score 12; DB 11; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 82 GCCACCATGGCC 93

RESULT 11

LOCUS AA863181/c 124 bp mRNA EST 29-APR-1998

DEFINITION Oq91f04.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1455679 3', similar to TR:Q94263 Q94263 SIMILAR TO THE RAS GENE FAMILY. ;, mRNA sequence.

ACCESSION AA863181

VERSION AA863181.1 GI:2955660

KEYWORDS EST.

SOURCE human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124)
NCI-CGAP http://www.ncbi.nlm.nih.gov/hicogap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 1333 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 1.

FEATURES
source
1. .124
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Kid5"

```

/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
AAGTGGAGAAATTCGCGCGCAATATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
26 a 25 c 53 g 20 t
BASE COUNT
ORIGIN

Query Match 100.0%; Score 12; DB 10; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 25 GCCACCATGGCC 14

RESULT 12
BF092125
LOCUS
DEFINITION RCO-TN0079-150900-026-h11 TN0079 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF092125
VERSION BF092125.1 GI:10897835
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml.pl?tl=st2-RC0-TN0079-150
900-026-h11&t3=2000-09-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 31
High quality sequence stop: 124.
Location/Qualifiers
1. .124
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TN0079"
/dev_stage="Adult"
/note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under

```

```

low stringency conditions."
30 a 26 c 39 g 29 t
BASE COUNT
ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 45 GCCACCATGGCC 56

RESULT 13
BG511069
LOCUS
DEFINITION sac65006.y1 Gm-cl072 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl072-179 5', similar to SW:NU02_SOLTU P80268 NADH-UBIQUINONE
OXIDOREDUCTASE 18 KD SUBUNIT ;, mRNA sequence.
ACCESSION BG511069
VERSION BG511069.1 GI:13481726
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 126)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com.
Location/Qualifiers
1. .126
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl072-179"
/clone_lib="Gm-cl072"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/note="vector: pbluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar PI
567374 is partially resistant to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a

```

poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Steve Clough (Lila Vodkin lab, University of Illinois).

BASE COUNT 28 a 34 c 34 g 30 t
 ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 126;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
 |||||
 Db 4 GCCACCATGGCC 15

RESULT 14
 AA214420
 LOCUS AA214420 127 bp mRNA EST 31-JAN-1997
 DEFINITION zq91h04.r1 StrataGene hNT neuron (#937233) Homo sapiens cDNA clone IMAGE:649399 5' similar to contains Alu repetitive element;; mRNA sequence.
 ACCESSION AA214420
 VERSION AA214420.1 GI:1813037
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 127)
 AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,B., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,M., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478

CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 116.
 Location/Qualifiers

FEATURES
 source
 1. 127
 /organism="Homo sapiens"
 /db_xref="GDB:5278679"
 /db_xref="taxon:9606"
 /clone="IMAGE:649399"
 /clone_lib="StrataGene hNT neuron (#937233)"
 /dev_stage="hNT neurons"
 /lab_host="SOLR (kanamycin resistant)"
 /note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Differentiated, post mitotic hNT neurons. Average insert size: 1.5 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT 18 a 38 c 31 g 36 t 4 others

ORIGIN

Query Match 100.0%; Score 12; DB 10; Length 127;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
 |||||
 Db 103 GCCACCATGGCC 114

RESULT 15
 AA716853
 LOCUS AA716853 130 bp mRNA EST 29-DEC-1997
 DEFINITION vu62c06.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone IMAGE:1195978 5', mRNA sequence.

ACCESSION AA716853
 VERSION AA716853.1 GI:2729127
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 130)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:643074

Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 79.
 Location/Qualifiers

FEATURES
 source
 1. 130
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1195978"
 /clone_lib="Soares_mammary_gland_NbMMG"
 /sex="male"
 /tissue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Organ: mammary gland; Vector: pT7u3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(df) primer [5'
 TGTTCACCAATCTGAAGTGGGCGCGCGCAATGTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7u3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 31 a 26 c 29 g 44 t
 ORIGIN

Query Match 100.0%; Score 12; DB 10; Length 130;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
| | | | | | | | | |
Db 17 GCCACCATGGCC 28

Search completed: January 17, 2002, 08:54:22
Job time: 7134 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2002, 11:49:54 ; Search time 10436.3 Seconds
(without alignments)
2445.411 Million cell updates/sec

Title: US-09-242-202A-27
Perfect score: 1547
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTCGCACAGCCTCTCCACACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo_hum.*
- 31: em.htgo_inv.*
- 32: em.htgo_rod.*
- 33: em.htg_hum.*
- 34: em.htg_inv.*
- 35: em.htg_rod.*
- 36: em.htg_other.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	765	49.5	885	12	SYNPIAN7V	L08875 PIAN7 cloni
2	756.2	48.9	902	12	ATPIVX	V00089 Cloning vec
3	756.2	48.9	902	12	PIVIX	X14353 plasmid pi-
4	756.2	48.9	902	12	SYNPIVXV	L08918 pi-VX cloni
5	755	48.8	2932	6	I07209	I07209 Sequence 3
6	556.2	42.4	3392	12	AB009864	AB009864 Expressio
7	607.4	39.3	7252	6	AX003206	AX003206 Sequence
8	605.8	39.2	2538	12	ASPMIN1	Z50148 Artificial
9	604.4	39.1	5865	6	A60212	A60212 Sequence 8
10	604.4	39.1	5865	6	AR122288	AR122288 Sequence
11	604.4	39.1	6028	6	A60209	A60209 Sequence 5
12	604.4	39.1	6028	6	AR122285	AR122285 Sequence
13	604.4	39.1	6061	6	A60210	A60210 Sequence 6
14	604.4	39.1	6061	6	AR122286	AR122286 Sequence
15	604.4	39.1	6312	6	A60211	A60211 Sequence 7
16	604.4	39.1	6312	6	AR122287	AR122287 Sequence
17	604.2	39.1	11795	6	AX027785	AX027785 Sequence
18	604.2	39.1	13254	6	AR038307	AR038307 Sequence
19	604.2	39.1	13254	6	AR038321	AR038321 Sequence
20	604.2	39.1	13254	6	I58596	I58596 Sequence 15
21	604.2	39.1	13254	6	I58610	I58610 Sequence 17
22	603	39.0	1905	6	AR027070	AR027070 Sequence
23	603	39.0	1905	6	I86203	I86203 Sequence 9
24	603	39.0	1968	12	AF087567	AF087567 Cloning v
25	603	39.0	1969	12	AF086840	AF086840 Cloning v
26	603	39.0	1969	12	AF086841	AF086841 Cloning v
27	603	39.0	1969	12	AF086842	AF086842 Cloning v
28	603	39.0	1969	12	AF086843	AF086843 Cloning v
29	603	39.0	1969	12	AF086844	AF086844 Cloning v
30	603	39.0	1969	12	AF086845	AF086845 Cloning v
31	603	39.0	1969	12	AF086846	AF086846 Cloning v
32	603	39.0	1969	12	AF086847	AF086847 Cloning v
33	603	39.0	1969	12	AF086848	AF086848 Cloning v
34	603	39.0	1969	12	AF086849	AF086849 Cloning v
35	603	39.0	1969	12	AF086850	AF086850 Cloning v
36	603	39.0	1969	12	AF086851	AF086851 Cloning v
37	603	39.0	1969	12	AF086852	AF086852 Cloning v
38	603	39.0	1969	12	AF086853	AF086853 Cloning v
39	603	39.0	1969	12	AF087562	AF087562 Cloning v
40	603	39.0	1969	12	AF087563	AF087563 Cloning v
41	603	39.0	1969	12	AF087564	AF087564 Cloning v
42	603	39.0	1969	12	AF087565	AF087565 Cloning v
43	603	39.0	1969	12	AF087566	AF087566 Cloning v
44	603	39.0	1969	12	AF087664	AF087664 Cloning v
45	603	39.0	1969	12	AF087665	AF087665 Cloning v

ALIGNMENTS

RESULT 1	SYNPIAN7V	885 bp	DNA	circular	SYN	26-JUL-1993
LOCUS	SYNPIAN7V	PIAN7 cloning vector.				
DEFINITION	L08875					
ACCESSION	L08875.1	GI:310776				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
COMMENT						

Synthetic construct DNA.
synthetic construct
artificial sequence.
1 (bases 1 to 885)
Gilbert.W.
Obtained from VecBase 3.0
Unpublished (1991)
These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program. PIAN7 -
Cloning vector
ENTRY PIAN7
TITLE PIAN7 - Cloning vector
#TYPE DNA CIRCULAR
DATE 12-SEP-1986

#sequence	16-DEC-1986
ACCESSION	VB0066
SOURCE	artificial
REFERENCE	
number 1	
#citation	sequence information from New England Biolabs COMMENT
Obtained 12-SEP-1986 from New England Biolabs	
by magnetic tape	
Revised 16-DEC-1986 by F. Pfeiffer:	
449/450 'AP' to 'TA' to match revised	
PIA7 is thought to replace PIVX.	
COMMENT	
The polylinker of PIA7 contains additional BglII and XbaI	
sites within the M13mp8/pUC8 polylinker.	
KEYWORDS	
CROSSREFERENCE	
#parent	
VecBase(3):PIVX	
#parent	
VecBase(3):pBR322, VecSource(3):Poly8, GenBank(50):EcoTgy	
PARENT	
Features of PIA7 (885 bp)	
residue	source
1- 202 202- 1 (c) Tyr-tRNA synthetic (GenBank(50):EcoTgy)	
198- 225 1- 28 part 1 of pUC8/M13mp8-polylinker	
234- 239 31- 36 part 2 of pUC8/M13mp8-polylinker	
240- 880 2522-3162 pBR322	
Conflict (cfl) and Mutations (mut): none	
FEATURE	
POLYLINKER	ECORI-SmaI-BamHI-Sali-PstI-BglII-XbaI-HindIII SELECTION
#suppressor	Supf
SUMMARY	PIA7 #length 885 #checksum 9656.
Location/Qualifiers	
i..885	
/organism="synthetic construct"	
/db_xref="taxon:32630"	
BASE COUNT	193 a 239 c 236 g 217 t
ORIGIN	
Query Match	49.5%; Score 765; DB 12; Length 885;
Best Local Similarity	95.9%; Pred. No. 6.3e-207;
Matches	805; Conservative 0; Mismatches 10; Indels 24; Gaps 1;
QY	478 TTTCGGACCTTTCAAGTGTATGTTGGGGGGAAGATTCGAACTTCGAACTCGATGAC 537
DB	1 TTTCGGACCTTTTCAAGTGTATGTTGGGGGGAAGATTCGAACTTCGAACTCGATGAC 60
QY	538 GGCAGATTAGAGTCGTCCCTTTGGCGCGTCGGGAACCCACACCGGGTAATGCTTTT 597
DB	61 GGCAGATTAGAGTCGTCCCTTTGGCGCGTCGGGAACCCACACCGGGTAATGCTTTT 120
QY	598 ACTGGCCCTGCTCCTTATCGGAAGCGGGGCGATCATATCAATATCAGCGCGCTGTAA 657
DB	121 ACTGGCCCTGCTCCTTATCGGAAGCGGGGCGATCATATCAATATCAGCGCGCTGTAA 180
QY	658 AGTGTACGTTGAGAAGATTC-----CTGCAGCCCGCGC 693
DB	181 AGTGTACGTTGAGAAGATTC-----CTGCAGCTCTTGAAGCTTG 240
QY	694 CGTTGCTGGCGTTTTCATAGCTCCGCCCCCTGACGAGCATCACAAAATCAGCGCT 753
DB	241 CGTTGCTGGCGTTTTCATAGCTCCGCCCCCTGACGAGCATCACAAAATCAGCGCT 300
QY	754 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGGTTTCCCTCGGAA 813
DB	301 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGCGGTTTCCCTCGGAA 360
QY	814 GCTCCCTCGTCCGCTCTCTGTTCACACCTGCCGCTTACCGGATACCTGTCGCGCTTC 873
DB	361 GCTCCCTCGTCCGCTCTCTGTTCACACCTGCCGCTTACCGGATACCTGTCGCGCTTC 420
QY	874 TCCTTCGGGAAGCGTGGCGCTTTCATAGCTCAGCGCTGTAGGTATCTCAGTTCGGTGT 933

```
QY 531 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACACGGGTAA 590
|
|
|
Db 169 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACACGGGTAA 228
|
|
|
QY 591 TGTCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGC 650
|
|
|
Db 229 TGTCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGC 288
|
|
|
QY 651 GCTGTAAGTGTACGTTGAGAAAGATTCCTGCACGCCCGCGTGTGCGGCTTTTTC 710
|
|
|
Db 289 GCTGTAAGTGTACGTTGAGAAAGATTCCTGCACGCCCGCGTGTGCGGCTTTTTC 337
|
|
|
QY 711 CATAGGCTCGGCCCTGACAGCATCAAAAAATCGAGCTCAAGTCAGAGTGGCGA 770
|
|
|
Db 338 CATAGGCTCGGCCCTGACAGCATCAAAAAATCGAGCTCAAGTCAGAGTGGCGA 397
|
|
|
QY 771 AACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGCTCT 830
|
|
|
Db 398 AACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGCTCT 457
|
|
|
QY 831 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTTCCCTTCGGGAAGCGTG 890
|
|
|
Db 458 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTTCCCTTCGGGAAGCGTG 517
|
|
|
QY 891 GCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAG 950
|
|
|
Db 518 GCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAG 577
|
|
|
QY 951 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTTACGCGCTTATCCGGTAAC 1010
|
|
|
Db 578 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTTACGCGCTTATCCGGTAAC 637
|
|
|
QY 1011 CGTCTGAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGCGCAGCCACTGGTAAC 1070
|
|
|
Db 638 CGTCTGAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGCGCAGCCACTGGTAAC 697
|
|
|
QY 1071 AGGATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGTGCGCTAAC 1130
|
|
|
Db 698 AGGATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGTGCGCTAAC 757
|
|
|
QY 1131 TAGCGCTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 1190
|
|
|
Db 758 TAGCGCTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 816
|
|
|
QY 1191 GAAAAGAGTGTGCTAGCTTTCATCGGCAAAACAAACCCGCTGGTACGCGTGGTTTT 1250
|
|
|
Db 817 GAAAAGAGTGTGCTAGCTTTCATCGGCAAAACAAACCCGCTGGTACGCGTGGTTTT 876
|
|
|
QY 1251 TTTGTTTGAAGCAGCAGATTACGC 1275
|
|
|
Db 877 TTTGTTTGAAGCAGCAGATTACGC 901
|
|
|
RESULT 3
PPIVX
LOCUS PPIVX 902 bp DNA SYN 07-FEB-1990
DEFINITION Plasmid pi-vx DNA.
ACCESSION X14353
VERSION X14353.1 GI:58268
KEYWORDS artificial sequence; origin of replication; plasmid; polylinker;
suppressor transfer RNA.
SOURCE
ORGANISM Plasmid synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 902)
AUTHORS Seed,B.
TITLE Purification of genomic sequences from bacteriophage libraries by
recombination and selection in vivo
JOURNAL Nucleic Acids Res. 11 (8), 2427-2445 (1983)
MEDLINE 83220795
FEATURES
Location/Qualifiers
1..902
source
```

```
misc_feature 1..109 /note="polylinker"
trna 110..212 /note="suppressor trna"
misc_feature 213..902 /note="origin of replication"
BASE COUNT 192 a 249 c 238 g 223 t
ORIGIN
Query Match 48.9%; Score 756.2; DB 12; Length 902;
Best Local Similarity 98.1%; Pred. No. 2e-204;
Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
QY 471 CGAATCTTTTCGAGCTTTTGAAGTGTGTTGGGGGAAGGATTGGAACCTTCGAAGT 530
|
|
|
Db 109 CGAATCTTTTCGAGCTTTTGAAGTGTGTTGGGGGAAGGATTGGAACCTTCGAAGT 168
|
|
|
QY 531 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACACGGGTAA 590
|
|
|
Db 169 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACACGGGTAA 228
|
|
|
QY 591 TGTCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGC 650
|
|
|
Db 229 TGTCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGC 288
|
|
|
QY 651 GCTGTAAAGTGTACGTTGAGAAAGATTCCTGCAGCCCGCGGTTGCTGGCGTTTTTC 710
|
|
|
Db 289 GCTGTAAAGTGTACGTTGAGAAAGATTCCTGCAGCCCGCGGTTGCTGGCGTTTTTC 337
|
|
|
QY 711 CATAGGCTCGGCCCTGACAGCATCAAAAAATCGAGCTCAAGTCAGAGTGGCGA 770
|
|
|
Db 338 CATAGGCTCGGCCCTGACAGCATCAAAAAATCGAGCTCAAGTCAGAGTGGCGA 397
|
|
|
QY 771 AACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGCTCT 457
|
|
|
QY 831 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTTCCCTTCGGGAAGCGTG 890
|
|
|
Db 458 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTTCCCTTCGGGAAGCGTG 517
|
|
|
QY 891 GCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAG 950
|
|
|
Db 518 GCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAG 577
|
|
|
QY 951 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTTACGCGCTTATCCGGTAAC 1010
|
|
|
Db 578 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTTACGCGCTTATCCGGTAAC 637
|
|
|
QY 1011 CGTCTGAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGCGCAGCCACTGGTAAC 1070
|
|
|
Db 638 CGTCTGAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGCGCAGCCACTGGTAAC 697
|
|
|
QY 1071 AGGATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGTGCGCTAAC 1130
|
|
|
Db 698 AGGATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTCTTGAAGTGTGCGCTAAC 757
|
|
|
QY 1131 TAGCGCTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 1190
|
|
|
Db 758 TAGCGCTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 816
|
|
|
QY 1191 GAAAAGAGTGTGCTAGCTTTCATCGGCAAAACAAACCCGCTGGTACGCGTGGTTTT 1250
|
|
|
Db 817 GAAAAGAGTGTGCTAGCTTTCATCGGCAAAACAAACCCGCTGGTACGCGTGGTTTT 876
|
|
|
QY 1251 TTTGTTTGAAGCAGCAGATTACGC 1275
|
|
|
Db 877 TTTGTTTGAAGCAGCAGATTACGC 901
|
|
|
```

RESULT 4
 SYNPIVXV 902 bp DNA circular SYN 26-JUL-1993
 Locus
 DEFINITION Pi-VX cloning vector used for screening bacteriophage lambda.
 ACCESSION L08918
 VERSION L08918.1 GI:310786
 KEYWORDS
 SOURCE Synthetic construct DNA.
 ORGANISM
 REFERENCE
 AUTHORS
 JOURNAL
 COMMENT
 TITLE Pi-VX - Cloning vector used for screening bacteriophage lambda
 DATE 17-DEC-1982
 ACCESSION V0010
 SOURCE artificial
 COLLECTION ATCC 39083
 REFERENCE
 #number 1
 #authors Seed B.
 #journal Nucl. Acids Res. (1983) 11: 2427-2445
 #title
 Purification of genomic sequences from bacteriophage libraries and selection in vivo.
 #comment bases 1-902
 #number 2
 #authors Maniatis T., Fritsch E.F., Sambrook J.
 #book (IN) Maniatis T., Fritsch E.F., Sambrook J.;
 MOLECULAR CLONING: 353-361;
 COLD SPRING HARBOR LABORATORY (1982).
 #comment BASES 1 TO 902
 REFERENCE 3
 #authors Backman K., Betlach M., Boyer H.W., Yanofsky S.
 #journal Cold Spring Harb Symp Quant Biol 43, 69-76 (1980)
 #title
 Genetic and Physical Studies on the replication of ColE1-type plasmids
 #comment
 the sequence of the PIVX origin is from pBR345 as published in this paper
 COMMENT
 The origin of replication is derived from pMB1 via pBR345 and sequence given for pBR345 differs at one position from that of pBR322. This may be a sequencing error.
 COMMENT
 The ATCC strain 39083 contains PIVX in combination with plasmid p3.
 COMMENT
 from GenBank
 USED FOR SCREENING BACTERIOPHAGE LAMBDA GENE LIBRARIES FOR SPECIFIC DNA SEQUENCES IN ESCHERICHIA COLI. PROBE SEQUENCES ARE INSERTED IN THE VECTOR PI-VX. CONTAINS A POLYLINKER, AND ORIGIN OF REPLICATION AND A TYROSINE AMBER-SUPPRESSOR GENE (SYNTHETIC SUPP GENE).
 KEY FROM TO DESCRIPTION
 ORGRL 600
 ORIGIN OF REPLICATION KEYWORDS
 suppressor gene \\
 polylinker \\
 origin of replication
 CROSSREFERENCE
 #identical
 GenBank(50):PIVX, EMBL(11):ATPIVX
 #parent

VecBase(3):pBR322, GenBank(50):EcoTgy, VecSource(3):OriMB1
 #offspring
 VecBase(3):PIAN7, VecBase(3):mWB2344
 PARENT
 Features pf PIVX (902 bp)
 residue source
 4- 37 1- 34 pBR322
 1- 115 1- 115 PIVX-Polylinker
 111- 317 207- 1 (c) Tyr-tRNA synthetic (GenBank(50):EcoTgy)
 320- 901 1- 583 OriMB1
 320- 901 2522-3104 pBR322
 Conflict (cfl) and Mutations (mut):
 PIVX source
 cfl 811-12 T.C TAC 492-94 OriMB1
 cfl 811-12 T.C TAC 3013-15 pBR322
 FEATURE
 Supf
 POLYLINKER EcoRI-ClaI-HindIII-XbaI-BglII-PstI-BamHI-EcoRI
 SELECTION
 #suppressor Supf
 SUMMARY PIVX #length 902 #checksum 9265.
 Location/Qualifiers
 1..902
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 192 a 250 c 237 g 223 t
 ORIGIN
 Query Match 48.9%; Score 756.2; DB 12; Length 902;
 Best Local Similarity 98.1%; Pred. No. 2e-204;
 Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
 QY 471 CGAATCTTCGGACTTTTGAAAGTGATGGTGGGGGAAGGATTCGAACCTTCGAAGT 530
 DB 109 CGAATCTTCGGACTTTTGAAAGTGATGGTGGGGGAAGGATTCGAACCTTCGAAGT 168
 QY 531 CGATGACGGCAGATTACAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA 590
 DB 169 CGATGACGGCAGATTACAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA 228
 QY 591 TCGTTTACTGGCTGCTGCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGGGCC 650
 DB 229 TCGTTTACTGGCTGCTGCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGGGCC 288
 QY 651 GCTGTAAAGTGTACGTTGAGAAAGATTCCTGCAGCCCGCGCTGCTGGCGTTTTC 710
 DB 289 GCTGTAAAGTGTACGTTGAGAAAGATTCCTGCAGCCCGCGCTGCTGGCGTTTTC 337
 QY 711 CATAGGCTCCGCCCTCGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGA 770
 DB 338 CATAGGCTCCGCCCTCGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGA 397
 QY 771 AACCCGACAGACTATAAAGATACCAAGCGGTTCCCTGGAAGCTCCCTCGTGGCTCT 830
 DB 398 AACCCGACAGACTATAAAGATACCAAGCGGTTCCCTGGAAGCTCCCTCGTGGCTCT 457
 QY 831 CCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 890
 DB 458 CCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 517
 QY 891 GCGTTTCTCAATGCTCACGCTAGTATCTCAGTTCGGTGTAGGTTCGCTCCAAG 950
 DB 518 GCGTTTCTCATAGCTCACGCTAGTATCTCAGTTCGGTGTAGGTTCGCTCCAAG 577
 QY 951 CTGGGCTGTGTGACCAACCCCGCTTACCGCCGACCGCTGCGCCCTTATCCGGTAACTAT 1010
 DB 578 CTGGGCTGTGTGACCAACCCCGCTTACCGCCGACCGCTGCGCCCTTATCCGGTAACTAT 637
 QY 1011 CGTCTTGTAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCGTAACT 1070
 DB 638 CGTCTTGTAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCGTAACT 697

```
QY 1071 AGGATTACAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1130
Db 698 AGGATTACAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 757
QY 1131 TACGCTACACTAGAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTC 1190
Db 758 TACGCTACACTAGAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTT-CCTTC 816
QY 1191 GGAAAAAGTTGGTAGTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTT 1250
Db 817 GGAAAAAGTTGGTAGTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTT 876
QY 1251 TTTGTTGCAAGCAGCAGATTACGC 1275
Db 877 TTTGTTGCAAGCAGCAGATTACGC 901

RESULT 5
107209/c
LOCUS 107209 2932 bp PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0330191.
ACCESSION 107209
VERSION 107209.1 GI:590060
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2932)
AUTHORS Seed,B.D.O.M.B., Allen,J., Aruffo,A., Camerini,D., Lauffer,L.D.,
Oquendo,C.P., Simmons,D., Stamenkovic,I. and Stengelin,S.D.
TITLE Rapid immunoselection cloning method
JOURNAL Patent: EP 0330191-A2 3 30-AUG-1989;
FEATURES Location/Qualifiers
source 1..2932
BASE COUNT 774 a 699 c 682 g 777 t
ORIGIN

Query Match 48.8%; Score 755; DB 6; Length 2932;
Best Local Similarity 97.6%; Pred. No. 4.4e-204;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAAATCTTTCGGACTTTTGAAAGTCATGCTGTGGGGAAGGATTCGACCTTCGAA 528
Db 804 AGCAGATTCTTTCGGACTTTTGAAAGTCATGCTGTGGGGAAGGATTCGACCTTCGAA 745
QY 529 GTCGATGACGCGAGATTTAGAGTCTGCTCCCTTTTGGCGCTCGGGAACCCACCACCGGT 588
Db 744 GTCGATGACGCGAGATTTAGAGTCTGCTCCCTTTTGGCGCTCGGGAACCCACCACCGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCAATGACGG 648
Db 684 AATGCTTTTACTGGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCAATGACGG 625
QY 649 CGCGTGAAGTGTACGTTGAGAAGAAATTCCTCAGCGCCCGCGTTCGTCGGCTTTT 708
Db 624 CGCGTGAAGTGTACGTTGAGAAGAA--CCGCGGTAATTCGCGGTTCGTCGGCTTTT 568
QY 709 TCCATAGGCTCCGCCCTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGC 768
Db 567 TCCATAGGCTCCGCCCTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGC 508
QY 769 GAAACCCGACAGGACTATAAGATACACAGGGCTTTCCCGCTGGAAGCTCCCTCGTGGCT 828
Db 507 GAAACCCGACAGGACTATAAGATACACAGGGCTTTCCCGCTGGAAGCTCCCTCGTGGCT 448
QY 829 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTCTCGCCCTTTCTCCCTTCGGGAAGCG 888
Db 447 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTCTCGCCCTTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTTCGGTTCGTTTCGCTCCA 948
Db 1135 TGGCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTTCGTTTCGCTCCA 194
```

```
Db 387 TGGCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTTCGTTTCGTTTCGCTCCA 328
QY 949 AGCTGGCTGTGTGCACGAACCCCGCTTACGCCGACGCTGCCTTATCCGGTAACT 1008
Db 327 AGCTGGCTGTGTGCACGAACCCCGCTTACGCCGACGCTGCCTTATCCGGTAACT 268
QY 1009 ATCGTCTTGAAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCACTGGTA 1068
Db 267 ATCGTCTTGAAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCACTGGTA 208
QY 1069 ACAGATTAGCAGACGAGGATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 1128
Db 207 ACAGATTAGCAGACGAGGATGTAGGGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 148
QY 1129 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 1188
Db 147 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT-CCT 89
QY 1189 TCGGAAAAAGTGTGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTT 1248
Db 88 TCGGAAAAAGTGTGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTT 29
QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 6
AB009864
LOCUS AB009864 3392 bp DNA circular SYN 26-DEC-1997
DEFINITION Expression vector pME18S-FL3, complete sequence.
ACCESSION AB009864
VERSION AB009864.1 GI:2723416
KEYWORDS expression vectors plasmid:pME18S-FL3 DNA.
SOURCE Plasmid vectors
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, The Institute of Medical Science, University of Tokyo,
Department of Virology; 4-6-1, Shirokanedai, Minatoku, Tokyo 108,
Japan (E-mail:sugano@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
REFERENCE 2 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE pME18S-FL3: a versatile expression vector
JOURNAL Published Only in Database (1997) In press
FEATURES Location/Qualifiers
source 1..3392
/organism="vectors"
/plasmid="pME18S-FL3"
/db_xref="taxon:29278"
promoter 1..638
polya_site 1229..1427
BASE COUNT 810 a 902 c 818 g 862 t
ORIGIN

Query Match 42.4%; Score 656.2; DB 12; Length 3392;
Best Local Similarity 70.8%; Pred. No. 6.6e-176;
Matches 990; Conservative 0; Mismatches 288; Indels 121; Gaps 4;

QY 15 TGGCGCGGATCTTTTATCACTGATAAGTTGGTGACATATTATGTTTATCACTGATAAG 74
Db 828 TGGCGCGGATCTTTTATCACTGATAAGTTGGTGACATATTATGTTTATCACTGATAAG 887
QY 75 TGTCAAGCATGACAAAGTTGCGAGCGAATACAGTATCGTCCCGCCCTGGACTGTGA 134
Db 888 TGTCAAGCATGACAAAGTTGCGAGCGAATACAGTATCGTCCCGCCCTAGACCTGTGA 947
QY 135 ACGAGTTCGGCTAGACGCTCTGACGACACGCAAACTGCGGGAACGGTTGGGGGTGCAC 194
```


unclassified.
REFERENCE
1 (bases 1 to 6028)
AUTHORS
Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE
EXPRESSION SYSTEMS
JOURNAL
Patent: WO 9708330-A 5 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)

[illegible]

REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 6 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES Location/Qualifiers
source
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGCGCGCTG 654
Db 3948 TTATCATGCTGGATCCAGATCTGGGCCCATCGCGCGGATCGATNNNACATGTGAG 4007
QY 655 TAAAGTGTACCTTGAGAAAGAAATTCCTGCAGCCCGCGGTTGCTGGCGTTTCCATA 714
Db 4008 CAAAAGCCAGCAAAAGCCAGAACCGTAAAGAGCCGCTTGTGCGGTTTCCATA 4067
QY 715 GGCTCCGCCCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 774
Db 4068 GGCTCCGCCCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 4127
QY 775 CGACAGGACTATAAAGTACCAGGCGTTTCCCTTGGAGCTCCCTCGCTCTCCTG 834
Db 4128 CGACAGGACTATAAAGTACCAGGCGTTTCCCTTGGAGCTCCCTCGCTCTCCTG 4187
QY 835 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTTCGCGGAAAGCGTGGCGC 894
Db 4188 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTTCGCGGAAAGCGTGGCGC 4247
QY 895 TTTCATGCTCAGCTGAGTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 954
Db 4248 TTTCATGCTCAGCTGAGTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 4307
QY 955 GCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACATATCGTC 1014
Db 4308 GCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACATATCGTC 4367
QY 1015 TTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 1074
Db 4368 TTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 4427
QY 1075 TTAGCAGAGGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAATCTAG 1134
Db 4428 TTAGCAGAGGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAATCTAG 4487
QY 1135 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4488 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4547
QY 1195 AAAGAGTTGGTAGTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTTG 1254
Db 4548 AAAGAGTTGGTAGTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTTG 4607
QY 1255 TTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCT 1292
Db 4608 TTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCT 4645

RESULT 14
AR122286
LOCUS AR122286 6061 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6165715.
ACCESSION AR122286
VERSION AR122286.1 GI:14106603
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.KatherineLevinge, Weiss,R.Anthony, Takeuchi,Y. and Cosset,F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 6 26-DEC-2000;
FEATURES Location/Qualifiers
source
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGCGCGCTG 654
Db 3948 TTATCATGCTGGATCCAGATCTGGGCCCATCGCGCGGATCGATNNNACATGTGAG 4007
QY 655 TAAAGTGTACCTTGAGAAAGAAATTCCTGCAGCCCGCGGTTGCTGGCGTTTCCATA 714
Db 4008 CAAAAGCCAGCAAAAGCCAGAACCGTAAAGAGCCGCTTGTGCGGTTTCCATA 4067
QY 715 GGCTCCGCCCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 774
Db 4068 GGCTCCGCCCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 4127
QY 775 CGACAGGACTATAAAGTACCAGGCGTTTCCCTTGGAGCTCCCTCGCTCTCCTG 834
Db 4128 CGACAGGACTATAAAGTACCAGGCGTTTCCCTTGGAGCTCCCTCGCTCTCCTG 4187
QY 835 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTTCGCGGAAAGCGTGGCGC 894
Db 4188 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTTCGCGGAAAGCGTGGCGC 4247
QY 895 TTTCATGCTCAGCTGAGTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 954
Db 4248 TTTCATGCTCAGCTGAGTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 4307
QY 955 GCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACATATCGTC 1014
Db 4308 GCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACATATCGTC 4367
QY 1015 TTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 1074
Db 4368 TTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 4427
QY 1075 TTAGCAGAGGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAATCTAG 1134
Db 4428 TTAGCAGAGGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAATCTAG 4487
QY 1135 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4488 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4547
QY 1195 AAAGAGTTGGTAGTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTTG 1254
Db 4548 AAAGAGTTGGTAGTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTTG 4607
QY 1255 TTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCT 1292
Db 4608 TTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCT 4645

RESULT 15
A60211
LOCUS A60211 6312 bp DNA circular PAT 06-MAR-1998
DEFINITION Sequence 7 from Patent WO9708330.
ACCESSION A60211

Job time: 17240 sec

```
VERSION A60211.1 GI:3715219
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 6312)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 7 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
source
1. .6312
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1575 a 1692 c 1561 g 1476 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6312;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAGCGGGGGCATATCAATGACGGCCGCTG 654
Db 4199 TTATCATGTCTGGATCCAGATCTGGGCCATCGCGCGGATCGATNNNACATGTGAG 4258
QY 655 TAAAGTGTACGTTGAGAAAGATTCTCGAGCCCGCGCTGCTGGCGTTTTCCTATA 714
Db 4259 CAAAGCCAGCAAAAGCCAGGACCGTAAAGAGCGCGTGTGCTGGCGTTTTCCTATA 4318
QY 715 GGCTCCGCCCCCTGAGGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774
Db 4319 GGCTCCGCCCCCTGAGGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 4378
QY 775 CGACAGACTATAAAGATACAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 834
Db 4379 CGACAGACTATAAAGATACAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 4438
QY 835 TTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGGGAAGGTGGCGC 894
Db 4439 TTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGGGAAGGTGGCGC 4498
QY 895 TTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTGCGTGTAGTGTGCTTCCAAAGCTGG 954
Db 4499 TTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTGCGTGTAGTGTGCTTCCAAAGCTGG 4558
QY 955 GCTGTGTGACGACACCCCGTTACGCCCCACCGCTCGCGCTTATCCGGTAACCTATCGTC 1014
Db 4559 GCTGTGTGACGACACCCCGTTACGCCCCACCGCTCGCGCTTATCCGGTAACCTATCGTC 4618
QY 1015 TTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCGCACCTGGTAACAGGA 1074
Db 4619 TTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCGCACCTGGTAACAGGA 4678
QY 1075 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCTTAACACG 1134
Db 4679 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCTTAACACG 4738
QY 1135 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 1194
Db 4739 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 4798
QY 1195 AAAGAGTTGTAGCTCTTGTATCCGGCAACAAACCACCGCTGTAGCGGTGTTTTTTT 1254
Db 4799 AAAGAGTTGTAGCTCTTGTATCCGGCAACAAACCACCGCTGTAGCGGTGTTTTTTT 4858
QY 1255 TTTGCAAGCAGCAGATTACCGCGAGAAAAAAGGATCT 1292
Db 4859 TTTGCAAGCAGCAGATTACCGCGAGAAAAAAGGATCT 4896
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:03:54 ; Search time 599.86 Seconds
(without alignments)
2210.986 Million cell updates/sec

Title: US-09-242-202A-27
Perfect score: 1547
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTGCACAGGCTCTCCACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1547	100.0	1547	19	Humanised vector p
2	1310.4	84.7	1807	19	Humanised vector p
3	1086.6	70.2	2308	19	Humanised vector p
C 4	755	48.8	2932	13	COS cell expressio
C 5	755	48.8	2932	17	Expression vector
C 6	755	48.8	2932	17	Vector plasmid piH
C 7	755	48.8	2932	19	Expression vector
C 8	755	48.8	2932	20	Expression vector
C 9	755	48.8	2932	20	Eukaryotic cell ex
C 10	755	48.8	2932	22	Expression vector
11	607.4	39.3	7252	20	Complete sequence

12	604.4	39.1	5865	18	AAT90592	Plasmid FBdelPRDSA
13	604.4	39.1	6028	18	AAT90689	Plasmid FBdelPASAF
14	604.4	39.1	6061	18	AAT90690	Plasmid FBdelPMOSA
15	604.4	39.1	6312	18	AAT90691	Plasmid FBdelPGASA
16	604.2	39.1	7073	18	AAT50962	TF8-569 CDR-graite
17	604.2	39.1	7864	18	AAT50963	TF8-569 CDR-graite
18	604.2	39.1	11795	21	AAT50962	Nucleotide sequenc
19	604.2	39.1	13254	17	AAT40915	Nucleotide sequenc
20	604.2	39.1	13254	21	AA332151	pEel2 Combo BM 12
21	604.2	39.1	13254	21	AA332151	Complete nucleotid
22	603	39.0	608	19	AAV21731	ColEI origin of re
23	603	39.0	2077	20	AAZ23771	Vector pASlib DNA.
24	603	39.0	2462	21	AA74638	Plasmid pSP72. Es
25	603	39.0	2577	14	AAQ36620	Expression plasmid
26	603	39.0	2704	20	AAZ06758	Vector pUC28 nucle
27	603	39.0	2927	17	AAZ29158	Plasmid pTRP. Syn
28	603	39.0	2927	17	AAZ31789	Plasmid pTRP. Syn
29	603	39.0	2939	16	AAQ87350	Plasmid BGINV. Sy
30	603	39.0	2939	19	AAV37293	pBGINV plasmid use
31	603	39.0	3003	11	AAQ05745	Plasmid P.L-mu-smc
32	603	39.0	3104	10	AA90296	DNA target sequenc
33	603	39.0	3130	19	AA46334	Universal reporter
34	603	39.0	3175	20	AAZ20066	DNA encoding gluta
35	603	39.0	3327	22	AAAC86263	Plasmid GHRH1-29YA
36	603	39.0	3331	20	AA84434	pBC66-01, containi
37	603	39.0	3331	21	AAZ99245	Nucleotide sequenc
38	603	39.0	3344	21	AA39495	Transgenic unc-119
39	603	39.0	3369	22	AAAC86255	pGHRH-4 41sk const
40	603	39.0	3403	9	AA80956	Plasmid pDS5/RBSII
41	603	39.0	3414	12	AAQ12785	pDS56/RBSII-2 sequ
42	603	39.0	3414	14	AAQ34609	Expression plasmid
43	603	39.0	3415	12	AAQ12784	pDS56/RBSII-1 sequ
44	603	39.0	3415	14	AAQ34608	Expression plasmid
45	603	39.0	3416	9	AA80958	Plasmid pDS56/RBSI

ALIGNMENTS

RESULT 1	
AAV21732	
ID	AAV21732 standard; cDNA; 1547 BP.
XX	
AC	AAV21732;
XX	
DT	17-AUG-1998 (first entry)
XX	
DE	Humanised vector pITL-A.
XX	
KW	Vector; vaccine; tumour; antigen; plasmid pITL-A; ds.
OS	Chimeric - Homo sapiens.
OS	Chimeric - Escherichia coli.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	472..680
FT	/*tag= c
FT	/note= "SupF gene"
FT	686..1292
FT	/*tag= d
FT	/note= "ColEI origin of replication"
FT	1311..1547
FT	/*tag= e
FT	/note= "RANTES promoter"
XX	
PN	WO9806863-A1.
XX	
PD	19-FEB-1998.
XX	
PF	14-AUG-1997; 97WO-US14306.
XX	
PR	14-AUG-1996; 96US-0023931.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Nelson EL, Nelson PJ;

PI WPI; 1998-159552/14.

XX Humanised polynucleotide vectors - comprising human derived promoter

PT and sequence acceptance site, used for the production of vaccines

XX Claim 14; Page 56-57; 125pp; English.

XX Plasmid pITL-A comprises a base vector for novel humanised
CC polynucleotide vectors. The vector in Escherichia coli DH
CC 10-beta/P3 is deposited as ATCC 98401. Novel vectors comprise a
CC human-derived promoter or mammalian homologue which is functional
CC in mammalian target tissue and cells and a sequence acceptance site
CC (see AAV21735-36) which accepts cDNA products from RI-PCR cloning.
CC They also contain minimal non-human components, such as a replication
CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
CC are necessary for production of the vector. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.

XX Sequence 1547 BP; 344 A; 408 C; 425 G; 370 T; 0 other;

Query Match 100.0%; Score 1547; DB 19; Length 1547;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCTGCGACCATGCGCGGATCTTTATCACTGATAAGTTGGTGACATATTATGTT 60
DB 1 ggtacctgcaccatgscggaattcttctacatgataagttgggacatatattggt 60
QY 61 TATCAGTGTAAAGTCTAAGCATGACAAAGTTGACGCGAATACAGTATCGTGCCGG 120
DB 61 tatcadtgataaagtctcaagcatgacaaagtgtgcagccgaatacacgtgacgtgccc 120
QY 121 CCTGTGACTGTTGAACGAGGTGCGGCTAGACGCTGACGACACGCAAACTGCGGAACG 180
DB 121 cctgtgactgttgaacgaggtgcggtgacgctgacgacgctgcagacacgcaactg9cggaacg 180
QY 181 GTTGGGGGTGACGACGCGCGCTTTTACTGGCCTTTACGAAACAGCGGCGCTTTAAGG 240
DB 181 gttgggggtgacgacgscggtcttacttgcaacttcagaaacagcggcgcttaagg 240
QY 241 GCCATATGTGAGTGAATGCTTACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
DB 241 gccatatgtgagtgtgactgacccagcggggatggggagacacctgtagtccagag 300
QY 301 CCCCGGGGACGACGCGCAATGCGCGTCTTCCCTGCGAGGATGAGTAGTGAGTCGCTC 360
DB 301 ccccggggacgacagcgcaatgcccgtcttccctgacgagatgagtgtgagtgccctc 360
QY 361 TCTGCGCCTTGAAGTTGCACTCCAGTCCGACCCAGCCTTGTCTTAATAAATAAATTAAGTT 420
DB 361 tcttgccttgaagtggcaactccagtcacccagcctgtcctaataaaataaatt 420
QY 421 GCATCATTTTGTCTGACTAGGTGCTCTATATATATTAAGCTTGATATCGAATTTCTTT 480
DB 421 gcatcatatttgtgactaggtgtctctatataatattataaagtgtgatatcgaaattcttt 480
QY 481 CGGACTTTTGAAGTATGATGGTGTGGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
DB 481 cggacttttgaagtgtggtggtggggaagatttcgaaccttcgaagtcgatgacgacg 540

QY 541 AGATTAGAGTCTGCTCCCTTTGGCGCTCGGAAACCCACACCGGTAATGCTTTTACT 600
DB 541 agatttagagctgctccctttggcgctcgggaaacccacacccggttaagtctttact 600
QY 601 GGCCTGCTCCCTTATTCGGGAAGCGGGGCGCATCATATCAATATGACGCGCGTGTAAAGT 660
DB 601 ggcctgctcccttattcgggaagcggggcgcatcatcaaatgacgcgcgcgtgtaagt 660
QY 661 GTTACGTTGAGAAAGAAATTCCTGCAGCGCGCGCTTGTGCTGCGTTTTCATAGGCTCC 720
DB 661 gttacgttgagaaagaattctcgcagccgcgcgtgtgctg9cggtttttccataagctcc 720
QY 721 GCGCCCTTACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACAG 780
DB 721 gccccctgaagcagcatcacaaaaatcgacgctcaagtgcagaggtggcgaaacccgacag 780
QY 781 GACTATAAAGATACGAGCGGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTGTTCCGA 840
DB 781 gactataaagataccagcgggtttccccctggaaagctccctcgtgcgctcctctgtttccga 840
QY 841 CCTGCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC 900
DB 841 cctcgcgcttaccggataacctgtccgcttctcccttcgggaagcgtggcgctttctc 900
QY 901 AATGCTACGCTGTAGGTATCTCAGTTCCGCTGAGTTCGTTCCGCTCCAGCTGGGCTGTG 960
DB 901 aatgctcacgctgtaggtatctcagttcgtgtagtgcgtcccaagctgggctgtg 960
QY 961 TGCAGAAACCCCGCTTACGCGCGGCTGAGCTGCTTATTCGCTTAACCTATCGTCTTTGAGT 1020
DB 961 tgcagaaaccccgcttaccgcccgcgcgcgtgcgcttaccggttaactatcgtcttgagt 1020
QY 1021 CCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA 1080
DB 1021 ccaacccgtaagacacgacttatcgcaactggcagcagcactggttaacaggtatagca 1080
QY 1081 GAGCAGAGTATGAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACACTACGCTACA 1140
DB 1081 gagcaggttatgtagcgggtgtacagagttcttgaagtgggtggcctaactacggtaca 1140
QY 1141 CTAGAGGACAGTATTTGTTGTTATCTGCGCTCTGTAAGCCAGTTACTTCGGAAGAAAGAG 1200
DB 1141 ctagaaggacagtatttggtaactcgcgtctcgtgaagccagttaccttcggaaaaagag 1200
QY 1201 TTGGTAGCTTTTATCCGCAACAAACACCCGCTGGTAGCGGTGGTGTGTTTGTGTTGCA 1260
DB 1201 ttggtagctcttgatccggcaacaaacccgcgtggtagcgggtgtttttttgttgca 1260
QY 1261 AGCAGCAGATTACGGCGCAAAAAGGATCTGGGGATCCGGGAGAGCTCACTCTAGATG 1320
DB 1261 agcagcagattacgcgcagaaaaaaggatctgggggatccggagagctcaactctatag 1320
QY 1321 AGAGACAGTGGGGAGACAGAGACTCGAATTTCCGAGAGCTATTTTCAAGTTTCTTTTTC 1380
DB 1321 agagagcagctggggagagacagagagactgaaattccggagcattatttcagtttcttcc 1380
QY 1381 CGTTTTGTGCAATTTTCACTTATGATACCGGCAATGCTTTGGTGTGCTTATTTGGAAGTCC 1440
DB 1381 cgtttgtgcaatttcaactatgataccggccaatgctggtgctatttttggaactcc 1440
QY 1441 CCTTAGGGATGCCCTCAACTGCGCCCTATTAAGGCCGAGCTGAGCTGCAGAGGATTCC 1500
DB 1441 ccttagggatgccctcaactggccctataaaggccagcctgagctgacagaggattcc 1500
QY 1501 TGCAGAGGATCAAGACAGCAGCTGGACCTCGCACACGCTCTCCACCA 1547
DB 1501 tgcagaggatcaagacagacgctggacctcgacagccctctccacca 1547

RESULT 2
AAV21733
ID AAV21733 standard; cDNA; 1807 BP.

||||| ctagaagacagtagtattgtgtatctgcgtctgctgaagccagttacccttcggaaaaagag 1200
Db 1141
||||| TTGGTAGCTCTTGATCGGCAAAACACCGCTGGTAGCGGTGGTTTGTGTTGCA 1260
Qy 1201
||||| ttggtagctctgacccggcaacacacccgctggtagcgggtgttttttttttttttga 1260
Db 1201
Qy 1261 AGCAGCAGATTACGCGCAAAAGAGATCTGGGGATCCGAGAGCTCAC 1312
Db 1261 agcagcagattacgcagaaaaaagatctggggatccgagagctccc 1312
Db 1261
RESULT 3
AAV21734
ID AAV21734 standard; cDNA: 2308 BP.
XX AC AAV21734;
XX AC AAV21734;
DT 17-AUG-1998 (first entry)
XX Humanised vector pITL-1 GFP.
DE Vector; vaccine; tumour; antigen; plasmid pITL-1 GFP;
KW green fluorescent protein; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Chimeric - Aequorea victoria.
OS Synthetic.
XX Key Location/Qualifiers
FH CDS 20..734
FT /*tag= a
FT /product= green fluorescent protein
FT 973..1181
FT /*tag= c
FT /note= "SupF gene"
FT 1191..1793
FT /*tag= d
FT /note= "ColE1 origin of replication"
FT 2063..2308
FT /*tag= e
FT /note= "RANTES promoter"
XX WO9806863-A1.
XX 19-FEB-1998.
XX 14-AUG-1997; 97WO-US14306.
XX 14-AUG-1996; 96US-0023931.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX Humanised polynucleotide vectors - comprising human derived promoter
PT and sequence acceptance site, used for the production of vaccines
XX Example 11; Page 57-58; 125pp; English.
XX plasmid pITL-1 GFP comprises base vector pITL-1 (see AAV21733) and
CC a humanised green fluorescent protein (GFP) reporter sequence
CC (see AAV21725). Novel humanised vectors comprise a human-derived
CC promoter or mammalian homologue which is functional in mammalian
CC target tissue and cells and a sequence acceptance site which
CC accepts cDNA products from RT-PCR cloning. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune

CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX
SQ Sequence 2308 BP; 571 A; 601 C; 604 G; 532 T; 0 other;

Query Match 70.2%; Score 1086.6; DB 19; Length 2308;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1089; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 220 GAACAAGCGCGGCTTAAGGCCATATGTTGAGTGGATGCTTGACCCAGCGGGAT 279
Db 721 gtacaagtgcgccttaaggccatattggtgagtggtgacctgaccccaaggcgggat 780
Qy 280 GGGGAGACCTGTAGTCAGAGCCCGGGCAGCACAGGCCAATGCCCTTCCCTGCG 339
Db 781 gggggagacctgtagtcagagcccccggcagacagccaatgccctccctccgc 840
Qy 340 AGGATGAGTAGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCCACAGCC 399
Db 841 aggatgagtagtgagtgccctctcctggccctggaagtggccactccagtggccaccagcc 900
Qy 400 TTGTCTTAATAAATTAAATTGTCATCTTCTGACTAGGTGCTCTATAATATTAT 459
Db 901 ttgtcctaataaattaaattgcatcattttgtctgactagtgctcctataattatt 960
Qy 460 AAGCTTGATATCGAATTCTTCGGACTTTTGAAGTGATGGTGGGGAAGGATTCCA 519
Db 961 aagcttgatatacgaattcttcgacttttgaagtgtggtggggaagattcga 1020
Qy 520 ACCTTCGAAGTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCC 579
Db 1021 accttcgaagtcgatgcagcagatttagagtcgtctcctttggcgtcggaacccc 1080
Qy 580 ACCACGGGTAAATGCTTTTACTGGCGTCTCCCTTATCGGGAAGCGGGCGCATATCA 639
Db 1081 accacgggtaagtctttactggcctgtcccttaccgggaagcggggcgatcatatca 1140
Qy 640 AATGACGGCGCGCTGTAAAGTGTGTACGTTGAGAAAGATTCTGACGCCCGCGGTTC 699
Db 1141 aatgacgcgcgcgtgtaagtgttaagttgagaagaattcctgcagccgcgcgttgc 1200
Qy 700 TGGCGTTTTTCATAGGCTCCGCCCTGAGAGCATCACAAAATTCAGCGTCAAGTC 759
Db 1201 tggcgttttccataggtccgcgccttcacacataagcagcagcagcagcagtc 1260
Qy 760 AGAGGTGGCGAAACCCGACGAGACTATAAGATACCAGCGCTTCCCGCTGGAAGCTCC 819
Db 1261 agagtggcgaacccgcagagactataaagataccaggcgtttccctctggaagctccc 1320
Qy 820 TCGTGCCTCTCTGTTCCGACCTGCCGCTTACCAGTACCTGTCGCGCTTTCCTCCCT 879
Db 1321 tcgtgcctctctctgttcgcagacctgcgcttaccgataacctgtccgctttccctt 1380
Qy 880 CGGGAAGCTGGCGCTTCTCAATGCTCAGCTGATAGGTATCTCAGTTGGGTAGTGTG 939
Db 1381 cgggaagcgtggcgctttctcaatgctcagctgtagggtatctcagttcgggtgagtg 1440
Qy 940 TTCGCTCCAAGCTGGGCTGTGTCAGCAACCCCGCTTCAGCCCGCAGCTGCGCTTAT 999
Db 1441 tcgcgtccaagctgggctgtgtgcaggaacccccccttcagcccccgcgcgtcgttat 1500
Qy 1000 CGGTAACCTATCTGCTTGTAGTCCAAACCCGCTAAGACACACTTATCGCCACTGGCAGCAG 1059
Db 1501 ccggtaaactatgctcttgatccaaaccccgtaagacacgacttatgcactggcagcag 1560
Qy 1060 CCAGCTGGTAACAGGATTACAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGT 1119
Db 1561 ccactggtaacaggattagcagagcaggtatgtaggcgtgtacagagtttcttgagtc 1620


```
QY 949 AGCTGGCTGTGTCAGCAACCCCGTTTCAGCCCGACCGCTCGGCTTATCCGGTAACCT 1008
Db 327 AGCTGGCTGTGTCAGCAACCCCGTTTCAGCCCGACCGCTCGGCTTATCCGGTAACCT 268
QY 1009 ATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
Db 267 ATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
QY 1069 ACAGGATTAGCAGGAGGAGTATGTAGGGGTGCTACAGAGTCTTTGAAGTGGTGGCCTA 1128
Db 207 ACAGGATTAGCAGGAGGAGTATGTAGGGGTGCTACAGAGTCTTTGAAGTGGTGGCCTA 148
QY 1129 ACTACGGCTACACTAGAAGACAGTATTTGGTATCTGCTGCTGAAGCCAGTTACCT 1188
Db 147 ACTACGGCTACACTAGAAGACAGTATTTGGTATCTGCTGCTGAAGCCAGTT-OCT 89
QY 1189 TCGGAAAAGAGTTGCTAGCTCTTGATCCGGCAACAAACACCACCGTGTAGGGTGGTT 1248
Db 88 TCGGAAAAGAGTTGCTAGCTCTTGATCCGGCAACAAACACCACCGTGTAGGGTGGTT 29
QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2
RESULT 5
ID AAT38557/c
XX AAT38557 standard; DNA; 2932 BP.
AC AAT38557;
XX
DT 21-JAN-1997 (first entry)
XX
DE Expression vector pIH3.
XX
KW CD19; cell surface antigen; T lymphocyte; immunoselection cloning;
KW vector; pIH3; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..589
FT /*tag= a
FT /*note= "bases derived from pMB1 origin (pBR322 ori)"
FT misc_feature 590..597
FT /*tag= b
FT /*function= SacI linker
FT misc_feature 598..799
FT /*tag= c
FT /*function= synthetic tyrosine suppressor trna gene
FT LTR 800..947
FT /*tag= d
FT /*function= ASV LTR fragment (PvuII to MluI)
FT enhancer 948..1500
FT /*tag= e
FT /*function= human cytomegalovirus AD169 enhancer
FT TATA_signal 1501..1650
FT /*tag= f
FT /*function= HIV TATA and tat-responsive elements
FT misc_feature 1651..1761
FT /*tag= g
FT /*function= pILNAN polylinker (HindIII to Xba)
FT misc_signal 1717..2569
FT /*tag= h
FT /*function= pSV splice and poly-Addition signals
FT rep_origin 2570..2917
FT /*tag= i
FT /*function= SV40 origin of replication (PvuII to HindIII)
FT misc_feature 2918..2922
FT /*tag= j
FT
```

```
FT
XX
PN EP739980-A2.
XX
PD 30-OCT-1996.
XX
PF 23-FEB-1989; 89EP-0103127.
XX
PR 25-FEB-1988; 88US-0160416.
XX
PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
XX
PI Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo C;
PI Seed B, Simmons D, Stamenkovic I, Stengelin S;
PI WPI; 1996-478744/48.
XX
XX
XX cDNA encoding CD19 cell surface antigen - useful for recombinant
XX antigen prodn. for diagnostic and therapeutic purposes
XX
XX Disclosure; Fig 1; 23pp; English.
XX
XX Expression vector pIH3 (AAT38557), deposited as ATCC 67634, can be
XX used in a novel immunoselection cloning method for cell surface
XX antigen genes, e.g. human CD19 (see also AAT38556). This involves
XX introducing cDNA fragments into the vector, transfecting mammalian
XX cells in culture, maintaining the cells under conditions that allow
XX antigen expression, exposing the cells to an antibody directed
XX against the antigen, and recovering positive cells by contact with
XX an immobilised antibody directed against the first antibody.
XX
XX Sequence 2932 BP; 773 A; 700 C; 682 G; 777 T; 0 other;
```

```
Query Match 48.8%; Score 755; DB 17; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.9e-219;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATTTCTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAA 528
Db 804 AGCAGATTTCTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAA 745
QY 529 GTCGATGACGCGCAGATTAGAGTCTCTCCCTTTGGCGCTCGGGAACCCACCACCGGT 588
Db 744 GTCGATGACGCGCAGATTAGAGTCTCTCCCTTTGGCGCTCGGGAACCCACCACCGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGGGGGCGCATCATCAATATGACGC 648
Db 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGGGGGCGCATCATCAATATGACGC 625
QY 649 CCGCTGTAAGTGTACGTTGAGAAAGAAATTCCTGCAGCCGCGCTGCTGCGGTTTT 708
Db 624 CCGCTGTAAGTGTACGTTGAGAAAGAAATTCCTGCAGCCGCGCTGCTGCGGTTTT 568
QY 709 TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGC 768
Db 567 TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGC 508
QY 769 GAAACCCGACAGGACTATAAAGATACCAAGGTTTCCCGCTGGAAGCTCCCTCTGCGCT 828
Db 507 GAAACCCGACAGGACTATAAAGATACCAAGGTTTCCCGCTGGAAGCTCCCTCTGCGCT 448
QY 829 CTCCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 888
Db 447 CTCCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTTCTCAATGCTCAGCTGAGGTATCTCAGTTTCGGTGTAGGTGCTGCTGCCA 948
Db 387 TGGCGCTTTCTCAATGCTCAGCTGAGGTATCTCAGTTTCGGTGTAGGTGCTGCTGCCA 328
QY 949 AGCTGGCTGTGTGCAGCAACCCCGTTTCAGCCCGACCGCTCGGCTTATCCGGTAACCT 1008
Db 327 AGCTGGCTGTGTGCAGCAACCCCGTTTCAGCCCGACCGCTCGGCTTATCCGGTAACCT 268
```

QY 1009 ATCGTCTTGAAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
 Db 267 ATCGTCTTGAAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
 QY 1069 ACAGGATTTAGCAGAGGAGGATATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCTA 1128
 Db 207 ACAGGATTTAGCAGAGGAGGATATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCTA 148
 QY 1129 ACTACGGCTACACTAGAGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTTACCT 1188
 Db 147 ACTACGGCTACACTAGAGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTT-CCT 89
 QY 1189 TCGGAAAGAGTTGGTAGTCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTT 1248
 Db 88 TCGGAAAGAGTTGGTAGTCTTTGATCCGGCAACAAACACCGCTGGTAGCGGTGGTT 29
 QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
 Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 6
 AAT14702/c
 ID AAT14702 standard; DNA; 2932 BP.
 XX
 AC AAT14702;
 XX
 DT 30-OCT-1996 (first entry)
 XX
 DE Vector plasmid pIH3.
 XX
 DE Cell surface antigen; cloning; immunoselection; immunotherapy;
 KW therapy; diagnosis; vector; pIH3; CD2; T-cell receptor; COS;
 KW T-lymphocyte; ss.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1..589 a
 FT /tag= b
 FT /note= "bases 1-589 are derived from the pMB1
 origin (pBR322 ori)"
 FT misc_feature 590..597
 FT /tag= b
 FT /note= "bases 590-597 are derived from the
 SacII linker"
 FT misc_feature 598..799
 FT /tag= c
 FT /note= "bases 598-799 are derived from the
 synthetic tyrosine suppressor tRNA
 gene (supF gene)"
 FT LTR 800..947
 FT /tag= d
 FT /note= "bases 800-947 are derived from a remnant
 of the ASV LTR fragment (PvuII-MluI)"
 FT enhancer 948..1500
 FT /tag= e
 FT /note= "bases 948-1500 are derived from the hCMV
 AD169 enhancer"
 FT misc_signal 1501..1650
 FT /tag= f
 FT /note= "bases 1501-1650 are derived from HIV TATA
 and tat-responsive elements"
 FT misc_feature 1651..1716
 FT /tag= g
 FT /note= "bases 1651-1716 are derived from the
 pLXAN polylinker (HindIII-Xba)"
 FT misc_signal 1717..2569
 FT /tag= h
 FT /note= "bases 1717-2569 are derived from pSV
 to splice and poly-Addition signals"
 FT misc_feature 2570..2917

FT /*tag= i
 FT /note= "bases 2570-2917 are derived from the SV40
 origin of replication (PvuII-HindIII)"
 FT misc_feature 2918..2922
 FT /tag= j
 FT /note= "bases 2918-2922 are derived from pIVX,
 remnant of R1 site from polylinker"
 XX
 PN US5506126-A.
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 88US-0160416.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 PR 18-OCT-1993; 93US-0139273.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Aruffo A, Seed B;
 XX
 DR WPI; 1996-200279/20.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 of diagnostic and therapeutic proteins
 XX
 PS Example 1; Fig 1A-B; 79pp; English.
 XX
 CC Vector pIH3 (AAT14702) is a COS cell expression vector constructed
 from pISV by inserting a synthetic transcription unit between the
 suppressor tRNA gene and the SV40 origin. This novel vector allows
 the generation of very large mammalian expression libraries and
 yields large amounts of protein in mammalian host cells, resulting
 in efficient selection. Efficiency of library construction is
 comparable to that achieved with phage lambda, but the clones
 generated are easier to manipulate. A novel immunoselection
 cloning method was used to clone genes (see also AAT14703-04 and
 AAT14706-26) for cell surface antigens of human lymphocytes.
 XX
 SQ Sequence 2932 BP; 774 A; 698 C; 683 G; 777 T; 0 other;

Query Match 48.88; Score 755; DB 17; Length 2932;
 Best Local Similarity 97.68; Pred. No. 3.9e-219;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATTTCTTCGGACTTTTGAAGTGATGGTGGGGAAGGATTCGAACCTTCGAA 528
 Db 804 AGCAGATTCCTTCGGACTTTTGAAGTGATGGTGGGGAAGGATTCGAACCTTCGAA 745
 QY 529 GTCGATGACGCGCAGATTTAGAGTCTGCTCCCTTTGGCGGTCGCGGAACCCACCACCGGT 588
 Db 744 GTCGATGACGCGCAGATTTAGAGTCTGCTCCCTTTGGCGGTCGCGGAACCCACCACCGGT 685
 QY 589 AATGCTTTTACTGGCCTGCTCCCTTATCGGGAAGCGGGGCGCATCATATCAATGACGCG 648
 Db 684 AATGCTTTTACTGGCCTGCTCCCTTATCGGGAAGCGGGGCGCATCATATCAATGACGCG 625
 QY 649 CCGCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGCAGCCCGCGCTGCTGCGGTTTT 708
 Db 624 CCGCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGCAGCCCGCGCTGCTGCGGTTTT 568
 QY 709 TCCATAGGCTCCGCGCCCTCCCTGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGC 768
 Db 567 TCCATAGGCTCCGCGCCCTCCCTGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGC 508
 QY 769 GAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGT 828
 Db 507 GAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGT 448

QY 829 CTCCTGTTCCGACCCCTCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACG 888
 Db 447 CTCCTGTTCCGACCCCTCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACG 388
 QY 889 TGGGCTTTCTCAATGCTCAGCTGAGGTATCTCAGTTCCGTTAGGTGCTGCTCCA 948
 Db 387 TGGGCTTTCTCATAGCTCAGCTGAGGTATCTCAGTTCCGTTAGGTGCTGCTCCA 328
 QY 949 AGCTGGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTGCGCTTATCCGGTAAC 1008
 Db 327 AGCTGGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTGCGCTTATCCGGTAAC 268
 QY 1009 ATGCTCTGTAGTCCAAACCCCGGTGAAGACACACTTATGCCACTGGCAGCCACTGGTA 1068
 Db 267 ATGCTCTGTAGTCCAAACCCCGGTGAAGACACACTTATGCCACTGGCAGCCACTGGTA 208
 QY 1069 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCCTA 1128
 Db 207 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCCTA 148
 QY 1129 ACTACGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCT 1188
 Db 147 ACTACGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTT-CT 89
 QY 1189 TCGGAAAGAGTTGGTACCTCTGTATCGCGCAACAAACCCAGCTGTGTAGCGGTGTT 1248
 Db 88 TCGGAAAGAGTTGGTACCTCTGTATCGCGCAACAAACCCAGCTGTGTAGCGGTGTT 29
 QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
 Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 7

ID AAV63441/C
 XX AAV63441 standard; DNA; 2932 BP.

AC AAV63441;

XX 07-JUN-1999 (first entry)

DE Expression vector pIH3.

KW Vector; pIH3; cell surface antigen; lymphocyte; human; cloning;
 KW COS; ss...

XX Chimeric - Escherichia coli.

OS Chimeric - Human cytomegalovirus.

OS Chimeric - Human immunodeficiency virus.

XX Chimeric - Rhesus macaque polyoma virus.

Key Location/Qualifiers

FT misc_feature

FT 1..589

FT /*tag= a

FT /*note= "from pMB1 origin (pBR322 ori)"

FT misc_feature

FT 590..597

FT /*tag= b

FT /*note= "Sall linker"

FT misc_feature

FT 598..799

FT /*tag= c

FT /*note= "from synthetic tyrosine suppressor tRNA

FT LTR

FT 800..947

FT /*tag= d

FT /*note= "ASV LTR fragment"

FT enhancer

FT 948..1500

FT /*tag= e

FT /*note= "human cytomegalovirus AD169 enhancer"

FT TATA_signal

FT 1501..1650

FT /*tag= f

FT /*note= "HIV TATA and tat-responsive elements"

FT misc_feature

FT 1651..1716

FT /*tag= g

FT misc_signal
 FT 1717..2569
 FT /*tag= h
 FT /*note= "SV40 small t antigen splice and early
 FT region polyA signals from pSV2"
 FT 2570..2917
 FT /*tag= i
 FT /*note= "SV40 origin of replication"
 FT 2918..2922
 FT /*tag= j
 FT /*note= "from pIVX, remnant of RI site from
 FT polylinker"

US5830731-A.

03-NOV-1998.

21-MAY-1997; 97US-0861205.

01-DEC-1992; 92US-0983647.

25-FEB-1988; 88US-0160416.

13-JUL-1989; 89US-0379076.

13-JUL-1990; 90US-0553759.

21-MAY-1997; 97US-0861205.

(GEO) GEN HOSPITAL CORP.

Aruffo A, Seed B;

WPI; 1998-609251/51.

New cloning vector and polylinker - based on existing sequences for efficient cloning and expression of mammalian cDNA(s), especially human lymphocyte antigenic sequences

Example 1; Fig 1A-B; 75pp; English.

This is the nucleotide sequence of expression vector pIH3, a COS cell expression vector constructed from pISV by inserting a synthetic transcription unit between the suppressor tRNA gene and the SV40 origin. The vector can be used in a novel method for cloning cDNAs from mammalian expression libraries. The method is based on transient expression of an antigen in eukaryotic cells and immunoselection of cells expressing the antigen by adhesion to an antibody-coated substrate. The method is useful for the isolation and molecular cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell. It has been used to clone genes encoding cell surface antigens associated with mammalian T lymphocyte antigens such as CD1-53, LFA-3, ICAM-1, FCRI, TLISA and Leu8 (see AAV63442-63 and AAW80440-55). The invention provides high efficiency expression vectors, such as pIH3, which allow the generation of very large mammalian expression libraries, and yield large amounts of protein in mammalian host cells, resulting in efficient selection. The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders of animals, including humans.

Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 48.8%; Score 755; DB 19; Length 2932;
 Best Local Similarity 97.6%; Pred. No. 3.9e-219;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTCGGACCTTTTGAAGTCATGCTGGTGGGGGAAGATTGCAACCTTCGAA 528

Db 804 AGCAGATTCTTTTCGGACCTTTTGAAGTCATGCTGGTGGGGGAAGATTGCAACCTTCGAA 745

QY 529 GTCGATGACGGCAGATTATAGAGTCTGCCCTTTGGCGGCTCGGGAACCCACACCGGTT 588

Db 744 GTCGATGACGGCAGATTATAGAGTCTGCCCTTTGGCGGCTCGGGAACCCACACCGGTT 685

```
QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGGC 648
DB 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGGC 625
QY 649 CCGCTGTAAGTGTACGTTGAGAAGAATTCCTGAGAGCCCGCGGTGCTGGCGTTT 708
DB 624 CCGCTGTAAGTGTACGTTGAGAAGA---CCGCGGTAAATTCGGCGTTGCTGGCGTTT 568
QY 709 TCCATAGGCTCCGCCCTCGAGGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGC 768
DB 567 TCCATAGGCTCCGCCCTCGAGGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGC 508
QY 769 GAAACCCGACGACTATAAAGTACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCT 828
DB 507 GAAACCCGACGACTATAAAGTACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCT 448
QY 829 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCTCGGGAAGC 888
DB 447 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCTCGGGAAGC 388
QY 889 TGGCGCTTCTCAATGCTCACGCTGAGGTATCTCAGTTCGGGTAGGTGCTGCTCCCA 948
DB 387 TGGCGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGGTAGGTGCTGCTCCCA 328
QY 949 AGCTGGCTGTGTGACGAACCCCGCTTTCAGCCCGACGCTGCGCTTATCCGGTAACT 1008
DB 327 AGCTGGCTGTGTGACGAACCCCGCTTTCAGCCCGACGCTGCGCTTATCCGGTAACT 268
QY 1009 ATCGCTTTCAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
DB 267 ATCGCTTTCAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
QY 1069 ACAGGATTACGAGCGAGGTATGTAGGCGGTCTACAGATTCCTGAAGTGTGGCCTA 1128
DB 207 ACAGGATTACGAGCGAGGTATGTAGGCGGTCTACAGATTCCTGAAGTGTGGCCTA 148
QY 1129 ACTACGCTTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 1188
DB 147 ACTACGCTTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTT-CCT 89
QY 1189 TCGGAAAAGAGTTGTAGCTCTTGATCCGCAAAACCAACCCGCTGGTGGTGGTT 1248
DB 88 TCGGAAAAGAGTTGTAGCTCTTGATCCGCAAAACCAACCCGCTGGTGGTGGTT 29
QY 1249 TTTTCTTTTTCAGCAGCAGATTACGC 1275
DB 28 TTTTCTTTTTCAGCAGCAGATTACGC 2

RESULT 8
AAV81199/c
ID AAV81199 standard; DNA; 2932 BP.
XX AC
XX AAV81199;
XX AC
DT 10-MAY-1999 (first entry)
XX DE
XX Expression vector piH3.
XX KW
XX CD40; cell surface antigen; human; vector; plasmid piH3;
XX KW
XX cDNA library; COS; ss.
XX OS
XX Chimeric - Escherichia coli.
XX OS
XX Chimeric - Human cytomegalovirus.
XX OS
XX Chimeric - Human immunodeficiency virus.
XX OS
XX Chimeric - Rhesus macaque polyoma virus.
XX FH
XX Key Location/Qualifiers
FT misc_feature 1..589
FT FT /*tag= a
FT FT /note= "from pMB1 origin (pBR322 ori)"
FT FT misc_feature 590..597
```

```
FT FT /*tag= b
FT FT /note= "Sali linker"
FT FT misc_feature 598..799
FT FT /*tag= c
FT FT /note= "from synthetic tyrosine suppressor tRNA
FT FT gene (supf gene)"
FT FT LTR 800..947
FT FT /*tag= d
FT FT /note= "ASV LTR fragment"
FT FT enhancer 948..1500
FT FT /*tag= e
FT FT /note= "human cytomegalovirus AD169 enhancer"
FT FT TATA_signal 1501..1650
FT FT /*tag= f
FT FT /note= "HIV TATA and tat-responsive elements"
FT FT misc_feature 1651..1716
FT FT /*tag= g
FT FT /note= "pILMXAN polylinker"
FT FT misc_signal 1717..2569
FT FT /*tag= h
FT FT /note= "SV40 small t antigen splice and early
FT FT region polyA signals from pSV2"
FT FT misc_feature 2570..2917
FT FT /*tag= i
FT FT /note= "SV40 origin of replication"
FT FT misc_feature 2918..2922
FT FT /*tag= j
FT FT /note= "piVX, remnant of R1 site from polylinker"
XX XX
XX XX us5849898-A.
XX XX 15-DEC-1998.
XX XX 07-JUN-1995; 95US-0485447.
XX XX 01-DEC-1992; 92US-0983647.
XX XX 25-FEB-1988; 88US-0160416.
XX XX 13-JUL-1989; 89US-0379076.
XX XX 23-MAR-1990; 90US-0498509.
XX XX 13-JUL-1990; 90US-0553759.
XX XX 07-JUN-1995; 95US-0485447.
XX XX (GEO ) GEN HOSPITAL CORP.
XX XX Allen J., Amiot M., Aruffo A., Camerini D., Lauffer L.;
XX XX Oquendo C., Seed B., Simmons D., Stamenkovic I., Stengelin S;
XX XX WPI; 1999-069813/06.
XX XX
XX XX cDNA encoding human CD40 antigen - useful for cloning cDNA encoding
XX XX cell surface antigens, constructing cDNA libraries, expression
XX XX vectors for expression in eukaryotic cells or their fragments
XX XX
XX XX Example 1; Fig 1A-B; 79pp; English.
XX XX
XX XX This is the nucleotide sequence of expression vector piH3, a COS
XX XX cell expression vector constructed from pISV by inserting a
XX XX synthetic transcription unit between the suppressor tRNA gene and
XX XX the SV40 origin. The vector can be used in a novel method for
XX XX cloning cDNAs from mammalian expression libraries. The method is
XX XX based on transient expression of an antigen in eukaryotic cells and
XX XX physical selection of cells expressing the antigen by adhesion to
XX XX an antibody-coated substrate, such as a culture dish. The method
XX XX is useful for the isolation and molecular cloning of any protein
XX XX which can be expressed and transported to the cell surface membrane
XX XX of a eukaryotic cell. The invention specifically provides the CD40
XX XX cDNA sequence (see AAV81198). It also provides high efficiency
XX XX expression vectors, such as piH3, which allow the generation of
XX XX very large mammalian expression libraries, and yield large amounts
XX XX of protein in mammalian host cells, resulting in efficient selection.
XX XX
XX XX Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;
```

Query Match 48.8%; Score 755; DB 20; Length 2932;
 Best Local Similarity 97.6%; Pred. No. 3.9e-219;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTTCGGAGCTTTTCAAGAGTGATGTGGGGGAGGATTCGAACCTTCGAA 528
 |||
 Db 804 AGCAGATCTTTTCGGAGCTTTTCAAGAGTGATGTGGGGGAGGATTCGAACCTTCGAA 745

QY 529 GTCCGATGAGCGGAGATTTAGAGTCTGCTCCCTTTTGGCGGCTGCGGAAACCCACACCGGT 588
 |||
 Db 744 GTCCGATGAGCGGAGATTTAGAGTCTGCTCCCTTTTGGCGGCTGCGGAAACCCACACCGGT 685

QY 589 AATGCTTTTACTGGCTCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAATGACGCG 648
 |||
 Db 684 AATGCTTTTACTGGCTCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAATGACGCG 625

QY 649 CCGCTGTAAAGTGTACCTTCGAGAAAGATTCCTTCGACCCCGCGGTCATCATATCAATGACGCG 708
 |||
 Db 624 CCGCTGTAAAGTGTACCTTCGAGAAAGATTCCTTCGACCCCGCGGTCATCATATCAATGACGCG 568

QY 709 TCCATAGGCTCGGCGCCCTCGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGC 768
 |||
 Db 567 TCCATAGGCTCGGCGCCCTCGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGC 508

QY 769 GAAACCCGACAGGACTATAAGATACACAGGCTTCCCTCGAAGTCCCTCGTGGCGT 828
 |||
 Db 507 GAAACCCGACAGGACTATAAGATACACAGGCTTCCCTCGAAGTCCCTCGTGGCGT 448

QY 829 CTCCTGTTCGACCTCGCGCTTACCGGATACCTTCGCGCTTCTCCCTTCGCGGAGCG 888
 |||
 Db 447 CTCCTGTTCGACCTCGCGCTTACCGGATACCTTCGCGCTTCTCCCTTCGCGGAGCG 388

QY 889 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTTCGCTCCA 948
 |||
 Db 387 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTTCGCTCCA 328

QY 949 AGTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACGCTGCGCTTATCCGTTAACT 1008
 |||
 Db 327 AGTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACGCTGCGCTTATCCGTTAACT 268

QY 1009 ATCGTCTTGAGTCCACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCTGTA 1068
 |||
 Db 267 ATCGTCTTGAGTCCACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCTGTA 208

QY 1069 ACAGGATTAGCAGGAGGTATGTAGCGGTGCTACAGCTTCTGAAGTGTGGCCTA 1128
 |||
 Db 207 ACAGGATTAGCAGGAGGTATGTAGCGGTGCTACAGCTTCTGAAGTGTGGCCTA 148

QY 1129 ACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACT 1188
 |||
 Db 147 ACTACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACT 89

QY 1189 TCGGAAAAGAGTTGCTAGCTCTTATCCGGCAACAAACACCCTGGTACCGGTGGTT 1248
 |||
 Db 88 TCGGAAAAGAGTTGCTAGCTCTTATCCGGCAACAAACACCCTGGTACCGGTGGTT 29

QY 1249 TTTTGTGTTGCAAGCAGGAGTATGCG 1275
 |||
 Db 28 TTTTGTGTTGCAAGCAGGAGTATGCG 2

RESULT 9
 AAA50577/c
 ID AAA50577 standard; DNA; 2932 BP.
 XX
 AC AAA50577;
 XX
 DT 19-DEC-2000 (first entry)
 XX
 DE Eukaryotic cell expression vector plasmid pIH3.
 XX
 KW Plasmid pIH3; vector; eukaryotic cell; COS; cell surface antigen;

immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 immune disorder; infection; asthma; immune-complex disease;
 ankyloidosis; multiple sclerosis; parasitic disease;
 immunoselection; panning; ds.
 XX
 OS Chimeric - Rhesus macaque polyoma virus.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Human cytomegalovirus.
 OS Chimeric - Human immunodeficiency virus.
 OS Chimeric - Synthetic.
 XX
 Key Location/Qualifiers
 misc_feature 1..589
 /tag= a
 /function= "replication origin"
 /note= "pBR322 ori from pMB1"
 590..597
 /tag= b
 /note= "derived from SacI linker"
 598..799
 /tag= c
 /gene= "supF"
 /note= "synthetic tyrosine suppressor tRNA gene"
 800..947
 /tag= d
 /note= "remnant of ASV LTR fragment"
 948..1500
 /tag= e
 /note= "human cytomegalovirus AD169 enhancer"
 1501..1650
 /tag= f
 /note= "HIV TATA and tat-responsive element"
 1651..1716
 /tag= g
 /note= "derived from pILNAN polylinker"
 1717..2569
 /tag= h
 /note= "derived from PSV to splice and poly-A site"
 2570..2917
 /tag= i
 /function= "replication origin"
 /note= "from SV40 origin of replication"
 2918..2922
 /tag= j
 /note= "from pIVX"
 US6111093-A.
 29-AUG-2000.
 XX
 28-OCT-1998; 98US-0181612.
 XX
 01-DEC-1992; 92US-0983647.
 25-FEB-1988; 88US-0160416.
 13-JUL-1989; 89US-0379076.
 23-MAR-1990; 90US-0498809.
 13-JUL-1990; 90US-0553759.
 XX
 (GEO) GEN HOSPITAL CORP.
 XX
 Stamenkovic I, Seed B;
 WPI; 2000-586382/55.
 XX
 Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 useful for immunodiagnosis and immunotherapy of immune-mediated
 infections or disorders, e.g. asthma, immune-complex disease, parasitic
 diseases
 XX
 Example 1; Fig 1A-B; 75pp; English.
 PS
 The present sequence is that of COS cell expression vector
 CC
 plasmid pIH3. The vector was constructed from pISV by inserting

CC a synthetic transcription unit between the suppressor tRNA gene and
 CC the SV40 origin. The transcription unit consisted of a chimeric
 CC promoter composed of human cytomegalovirus AD169 immediate early
 CC enhancer sequences fused to the HIV LTR -67 to +80 sequences.
 CC Immediately downstream from the LTR +80 was inserted a polylinker
 CC containing 2 BstXI sites separated by a 350 bp stuffer. The BstXI
 CC sites were flanked by XbaI sites, which could also be used to
 CC excise the insert. Downstream from the polylinker were placed the
 CC SV40 small t antigen splice and early region polyA signals derived
 CC from pSV2. The vector can be used to construct large mammalian
 CC cDNA expression libraries. A new method for cloning cDNA encoding
 CC cell surface antigens is based upon transient expression of antigen
 CC in eukaryotic (e.g. COS) cells and physical selection of cells
 CC expressing the antigen by adhesion to (panning on) an antibody
 CC coated substrate such as a culture dish. The purified cDNA and the
 CC protein that it encodes are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders in
 CC animals, including humans. These disorders include asthma,
 CC immune-complex disease, amyloidosis, parasitic diseases or multiple
 CC sclerosis. High efficiency cDNA expression vectors such as pIH3
 CC were designed for use in the novel methods of the invention,
 CC yielding large amounts of protein in mammalian host cells,
 CC resulting in efficient selection.
 XX
 SQ Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 48.8%; Score 755; DB 21; Length 2932;
 Best Local Similarity 97.6%; Pred. No. 3.9e-219;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTCTTCGGACCTTTTGAAGTGATGGTGGGGGAAGGATTGCAACCTTCGAA 528
 DB 804 ACCAGATTCTTCGGACCTTTTGAAGTGATGGTGGGGGAAGGATTGCAACCTTCGAA 745

QY 529 GTCGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACACGGGT 588
 DB 744 GTCGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACACGGGT 685

QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGGCGCATCATCAATGACGG 648
 DB 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGGCGCATCATCAATGACGG 625

QY 649 CCCTGCTTAAAGTGTAGCTTGAAGAAATTCCTGCACGCCCGCGCTGCTGCGCTTTT 708
 DB 624 CCCTGCTTAAAGTGTAGCTTGAAGAAATTCCTGCACGCCCGCGCTGCTGCGCTTTT 568

QY 709 TCCATAGGCTCCGCCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGC 768
 DB 567 TCCATAGGCTCCGCCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGC 508

QY 769 GAAACCCGACGACATTAAGATACGAGCGCTTCCCTTGAAGCTCCTCGTGGCGT 828
 DB 507 GAAACCCGACGACATTAAGATACGAGCGCTTCCCTTGAAGCTCCTCGTGGCGT 448

QY 829 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCG 888
 DB 447 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCG 388

QY 889 TGGCGCTTCTCAATGCTACAGCTGATAGGTATCTCAGTTCGGGTAGGTGCTGCTCCCA 948
 DB 387 TGGCGCTTCTCAATGCTACAGCTGATAGGTATCTCAGTTCGGGTAGGTGCTGCTCCCA 328

QY 949 AGCTGGGCTGTGTCAGGAACCCCGCTTTCAGCCCGACCGCTTTCGCGCTTATCCGTAAC 1008
 DB 327 AGCTGGGCTGTGTCAGGAACCCCGCTTTCAGCCCGACCGCTTTCGCGCTTATCCGTAAC 268

QY 1009 ATCGTCTTGAAGTCCAAACCCGGTGAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
 DB 267 ATCGTCTTGAAGTCCAAACCCGGTGAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTA 208

QY 1069 ACAGGATTAGCAGCGAGGATATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGCGCTA 1128

DB 207 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGCGCTA 148
 QY 1129 ACTACGGCTACACTAGAGGACAGCATTTTGGTATCTCGCTCTCTGAAGCCACTTACCT 1188
 DB 147 ACTACGGCTACACTAGAGGACAGCATTTTGGTATCTCGCTCTCTGAAGCCAGTT-CCT 89
 QY 1189 TCGGAAAAAGAGTTGGTAGCTCTTTGATCCCGGCAACAAACACCGCTGGTAGCGTGTT 1248
 DB 88 TCGGAAAAAGAGTTGGTAGCTCTTTGATCCCGGCAACAAACACCGCTGGTAGCGTGTT 29
 QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
 DB 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 10
 AAS03171/c
 ID AAS03171 standard; DNA; 2932 BP.
 XX
 AC AAS03171;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE Expression vector pIH3 DNA used to express human lymphocyte antigens.
 XX
 KW Human; Lymphocyte cell surface antigen; immune-mediated disease; asthma;
 KW infection; immune deficiency disorder; hypersensitivity; inflammation;
 KW systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;
 KW transplant rejection; cyclic; circular; mutant; piH3; ds.
 XX
 OS Chimeric - Avian sarcoma virus.
 OS Chimeric - Human cytomegalovirus.
 OS Chimeric - Human immunodeficiency virus.
 OS Chimeric - Rhesus macaque polyoma virus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1..589 /*tag= a
 FT /*note= "Derived from the pMB1 origin (pBR322 ori)"
 FT misc_feature 590..597 /*tag= b
 FT /*note= "Derived from the SacII linker (ACCGGT)"
 FT misc_feature 598..799 /*tag= c
 FT /*note= "Derived from the synthetic tRNA gene (supP)"
 FT misc_feature 800..947 /*tag= d
 FT /*note= "Derived from a remnant of the ASV LTR
 FT fragment (PvuII to MluI)"
 FT enhancer 948..1500 /*tag= e
 FT /*note= "Derived from the human cytomegalovirus
 FT AD169 enhancer"
 FT promoter 1501..1650 /*tag= f
 FT /*note= "Derived from the HIV TATA and tat-responsive
 FT elements"
 FT misc_feature 1651..1716 /*tag= g
 FT /*note= "Derived from the pILNAN polylinker
 FT (HindIII to Xba)"
 FT misc_signal 1717..2569 /*tag= h
 FT /*note= "Derived from the pSV2 splice and polyadenylation
 FT signals"
 FT misc_feature 2570..2917 /*tag= i
 FT /*note= "Derived from the SV40 origin of replication
 FT (PvuII to HindIII)"
 FT misc_feature 2918..2922 /*tag= j

QY 715 GGCTCCGCCCTGAGGACATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACC 774
 Db 3872 ggctcgcgcctcctgacgagcatcacaaaatcgagctcaagtcagagtgcggaacc 3931
 QY 775 CGACAGGACTATAAGATACCAAGCGGTTCCTCCCTGGAAGCTCCCTGCGCTCCCTG 834
 Db 3932 cgacagactataaagataaccaggctttccctcctggaagctccctcgtcgctcctcg 3991
 QY 835 TTCGAGCCCTGCGCTTACCGGATACGTGCGCCCTTCTCCCTTCGGAAGCGTGGCGC 894
 Db 3992 ttcgagccctgcgcttacggataactgtccgcctttctccctcctggaagcgtggcg 4051
 QY 895 TTCTCAATGCTCACGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCAACTGG 954
 Db 4052 ttctcaatgctcacgctgtaggtatctcagttcgtgtaggtcgttcctcaagctgg 4111
 QY 955 GCTGTGTCGACGAACCCCGTTCACCGCGACCGCTGCGCTGCGCTTATCCCGTACTATGCTC 1014
 Db 4112 gctgtgtgcagaaaccccgcttcagccgacgctgcgcttattccggttaactatcgtc 4171
 QY 1015 TTGAGTCCAAACCGGTAAACACGACTTATCGCACTGCGCAGCAGCCTGTAAACAGA 1074
 Db 4172 ttgagtcacacccggttaacacagactattccactggcagcagcactggttaacagga 4231
 QY 1075 TTAGCAGCGAGGTATGAGCGGTGTCACAGAGTTCCTTGAAGTGGTGGCCTAACTAGC 1134
 Db 4232 ttgacagagcgaggtatgtagcggtgctacagagttcttgagtggtggcctaactag 4291
 QY 1135 GCTACACTAGAAGGACGATTTGGTATCTGCGCTGCTGAAGCCAGTTACTCTCGGAA 1194
 Db 4292 gctacactagaaggacagatttggatctgctgctgctgctgctgctgctgctgctgctg 4351
 QY 1195 AAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACACACCGCTGCTGAGCGTGGTTTTTG 1254
 Db 4352 aaagagtgtgtagctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 4411
 QY 1255 TTTCGAACGACGAGTTACGCGCCACAAAGAGATCT 1292
 Db 4412 ttgcaagcagcagattacgcgcagaaaaaaggatct 4449

RESULT 13
 ID AAT90689
 XX AAT90689 standard; DNA; 6028 BP.
 AC AAT90689;
 XX
 DT 05-JAN-1998 (first entry)
 DE Plasmid FBDelPASAF coding sequence.
 KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
 KW moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
 KW gene therapy; ss.
 OS Synthetic.
 XX
 PN W09708330-A1.
 PD 06-MAR-1997.
 XX
 PF 23-AUG-1996; 96WO-GB02061.
 XX
 PR 23-AUG-1995; 95GB-0017263.
 XX
 PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 XX
 PI Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
 XX WPI; 1997-179287/16.
 DR
 XX
 PT Selectable retroviral packaging cell lines and expression constructs

PT - comprise selectable gene downstream of gene of interest, are
 PT selectable due to the in-efficiency associated with translation
 XX re-initiation
 PS Claim 13; Fig 9; 79pp; English.
 XX This sequence represents the recombinant expression plasmid FBDelPASAF.
 CC (in this case from the moloney murine leukaemia virus) and a selectable
 CC marker (SM). It is an example of a recombinant expression vector (REV) of
 CC the invention, used to create a packaging cell line. The REV's of the
 CC invention comprise a gene of interest (GOI) and a SM gene. The SM gene is
 CC arranged downstream of the GOI and a GOI associated stop codon is spaced
 CC from a start codon of the SM gene to ensure that the SM protein is
 CC expressed as a result of translation reinitiation. The cell lines are
 CC transformed with two REV's, both are replication deficient, one contains
 CC the viral gag-pol gene, the other the viral env gene. By using helper
 CC constructs, such as the REV's, which are directly selectable and which
 CC provide for high expression of the viral gene, high titre retroviral
 CC vectors may be obtained. The packaging cell lines are useful for gene
 CC therapy. Prior packaging cell lines using full length retroviral genomes
 CC as helper genomes were isolated by cotransfecting them with plasmids
 CC encoding selectable markers. However, the helper functions can be lost
 CC during the passages of the cells in culture and the current packaging
 CC systems provide limited titres of infectious retroviral vectors.
 CC Co-transfection with a plasmid encoding a SM does not directly select the
 CC best gag-pol-env-expressing cells. The new retroviral packaging cell
 CC lines overcome these problems.
 XX
 SQ Sequence 6028 BP; 1515 A; 1637 C; 1499 G; 1369 T; 8 other;

Query Match 39.1%; Score 604.4; DB 18; Length 6028;
 Best Local Similarity 91.4%; Pred. No. 3.7e-173;
 Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 595 TTTACTGCTTCCTTATCGGAGCGGGCGCATCATCAATGACGCCCGCTG 654
 Db 3915 ttatcatgtctggtatccagatctggcccatgcygcgcgagatgcatnnacatgtgag 3974
 QY 655 TAAAGTGTGAGTGAAGAATCTCTGACGCCCGCTTCTGCTGCGCTTTCCTCAT 714
 Db 3975 caaaagccagcaaaagccaggaacccgtaaaagcgctgtgctgctgttttcata 4034
 QY 715 GCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAACC 774
 Db 4035 ggctccgccccctgacgagcatcaaaaatcgacgtcaagtcagagtgcggaacc 4094
 QY 775 CGACAGGACTATAAGATACAGCGCTTTCCTCCCTGGAAGCTCCCTGCTCTCTG 834
 Db 4095 cgacagactataaagataaccagcgctttccctcctggaagctccctcgtcgtcctcg 4154
 QY 835 TTCGAGCCCTGCGCTTACCGGATACCTGTCGCTTTCCTTCCCTTCGGGAGCGTGGCGC 894
 Db 4155 ttcgagccctgcgcttaccggataacctgctcccttctcccttcgggaagcgtggcg 4214
 QY 895 TTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGCTGTAGTTCGCTCCAACTGG 954
 Db 4215 ttctcaatgctcacgctgtaggtatctcagttcgtgtaggtcgttcctcaagctgg 4274
 QY 955 GCTGTGTCAGCAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
 Db 4275 gctgtgtgcaaaccccccttcagcccgacgctgcttattccggttaactatcgtc 4334
 QY 1015 TTGAGTCCAAACCGGTAAACACGACTTATCGCACTGCGCAGCAGCCTGTAAACAGA 1074
 Db 4335 ttgagtcacacccggttaacacagactattccactggcagcagcactggttaacagga 4394
 QY 1075 TTAGCAGCGAGGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCCTAACTAGC 1134
 Db 4395 ttacagagcgaggtatgtagcggtgctacagagttcttgaagtggtggcctaactag 4454
 QY 1135 GCTACACTAGAAGGACAGATTTGGTATCTGCGCTGCTGGAAGCCAGTTACCTCGGAA 1194

||||| gctacactagaagacagtagtatttggtatctgcgtctctgctgaagccagttacacctcgaa 4514
Db 4455
QY 1195 AAGAGTTGCTAGCTCTTGATCCGGCAACAAACACCGCTGCTAGCGGTGTTTGTG 1254
Db 4515 aagagtggtgagctctgtatccgcgaacaaaccccgctgtagcgttggtttttg 4574
QY 1255 TTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db 4575 ttgcaagcagcagattacgcgcgagaaaaaaggatct 4612

RESULT 14
ID AAT90690 standard; DNA; 6061 BP.
XX
AC AAT90690;
XX
DT 05-JAN-1998 (first entry)
XX
DE Plasmid FBdelPMOSAF coding sequence.
XX
KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
KW gene therapy; ss.
XX
OS Synthetic.
XX
PN WO9708330-A1.
XX
PD 06-MAR-1997.
XX
XX 23-AUG-1996; 96WO-GB02061.
XX
PR 23-AUG-1995; 95GB-0017263.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
XX Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
XX WPI; 1997-179287/16.
XX
PT Selectable retroviral packaging cell lines and expression constructs
PT - comprise selectable gene downstream of gene of interest, are
PT selectable due to the in-efficiency associated with translation
PT re-initiation
XX
PS Claim 13; Fig 10; 79pp; English.
XX
CC This sequence represents the recombinant expression plasmid FBdelPASAF.
CC This sequence is a packaging-deficient construct having a viral env gene
CC (in this case from the moloney murine leukaemia virus) and a selectable
CC marker (SM). It is an example of a recombinant expression vector (REV) of
CC the invention, used to create a packaging cell line. The REV's of the
CC invention comprise a gene of interest (GOI) and a SM gene. The SM gene is
CC arranged downstream of the GOI and a GOI associated stop codon is spaced
CC from a start codon of the SM gene to ensure that the SM protein is
CC expressed as a result of translation reinitiation. The cell lines are
CC transformed with two REV's, both are replication deficient, one contains
CC the viral gag-pol gene, the other the viral env gene. By using helper
CC constructs, such as the REV's, which are directly selectable and which
CC provide for high expression of the viral gene, high titre retroviral
CC vectors may be obtained. The packaging cell lines are useful for gene
CC therapy. Prior packaging cell lines using full length retroviral genomes
CC as helper genomes were isolated by cotransfecting them with plasmids
CC encoding selectable markers. However, the helper functions can be lost
CC during the passages of the cells in culture and the current packaging
CC systems provide limited titres of infectious retroviral vectors.
CC Co-transfection with a plasmid encoding a SM does not directly select the
CC best gag-pol-env-expressing cells. The new retroviral packaging cell
CC lines overcome these problems.
XX
SQ Sequence 6061 BP; 1498 A; 1665 C; 1485 G; 1405 T; 8 other;

Query Match 39.1%; Score 604.4; DB 18; Length 6061;
Best Local Similarity 91.4%; Pred. No. 3.7e-173;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGGCTGCTCCCTTATCGGGAAGCGGGCGATCATCAATGACGGCGCGCTG 654
Db 3948 ttatcatgtctggtaccagatctggtcccatcgccgcgcatcatnnnnacatgtgag 4007
QY 655 TAAAGTGTTAGGTGAGAAAGATTCTGTCAGCCCGCGCTGCTGCGGCTTTTCCATA 714
Db 4008 caaagggccagcaaaaggccaggaacgcgtataaaggcgcgttgcgtgttttccata 4067
QY 715 GGCTCGCCGCCCTGACGAGCATCACAAAAATCGAGCTCAAGTCAGAGTGCGGAAACC 774
Db 4068 ggtccgcgccccctgacgagcatcacaaaaatcgacctcaagtcagaggtggcgaacc 4127
QY 775 CGACAGGACTATAAAGATACAGGCGGTTCCCGCTGGAAGCTCCCTGCTGCTCTCTG 834
Db 4128 cgacaggactataaagataccaggcgtttcccccctggaagctccctcgtcgctctcgt 4187
QY 835 TTCCGACCCCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCCGGGAGCGTGCGC 894
Db 4188 ttccgacctgcccgttaccggatacctgccgccttcccttcccttcgggaagcgtggcgc 4247
QY 895 TTTCTCAATGCTCAGCGCTGAGGTATCTCAGTTTCGGGTGAGGTCTGCTCCAAAGCTGG 954
Db 4248 ttctcaatgctcacgcgtgaggtatctcagttcgggtgaggtcgttcgctccaaagctgg 4307
QY 955 GCTGTGTGACGAAACCCCGCTTACGCCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
Db 4308 gctgtgtgacgaaccccccttcagccgcgacgcgtgccttatccgtaactatcgtc 4367
QY 1015 TTGAGTCCAAACCCGGTAAAGACACACTTATCGCCACTGGCAGCAGCCACTGTGTAACAGGA 1074
Db 4368 ttgagtcctcaacccggtaagacacgacttatcgccactggcagcagccactggtaacagga 4427
QY 1075 TTAGCAGAGCAGGTATGATAGCGGGTGTACAGAGTTCTTGAAGTGGTGCCTTAACACTACG 1134
Db 4428 ttgagcagagcagaggtatgtaggcgggtgtagcagagttcttgaggtggcgccctaacacg 4487
QY 1135 GCTACACTAGAAGGACACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTTCGGA 1194
Db 4488 gctacactagaagacagatttgggtatctgcgtctgctgaagccagttacacttcgaa 4547
QY 1195 AAGAGTTGTTAGTCTTTGATCCGGCAACAAACACCGCTGTTAGCGGTGTTTGTG 1254
Db 4548 aagagttggtagctcttgatccgcgcaaacacacccgctggtgtagcgttggttttttg 4607
QY 1255 TTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db 4608 ttgcaagcagcagattacgcgcgagaaaaaaggatct 4645

RESULT 15
ID AAT90691 standard; DNA; 6312 BP.
XX
AC AAT90691;
XX
DT 05-JAN-1998 (first entry)
XX
DE Plasmid FBdelPMOSAF coding sequence.
XX
KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
KW gene therapy; ss.
XX
OS Synthetic.
XX
PN WO9708330-A1.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:54:39 ; Search time 221.34 Seconds
(without alignments)
1582.911 Million cell updates/sec

Title: US-09-242-202a-27
Perfect score: 1547
Sequence: 1 GGTACTGTCACCATGGCGC.....CTGCACAGCCTCTCCACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2.6/ptodata/2/ina/5A_COMB.seq:*

2: /cgn2.6/ptodata/2/ina/5B_COMB.seq:*

3: /cgn2.6/ptodata/2/ina/6A_COMB.seq:*

4: /cgn2.6/ptodata/2/ina/6B_COMB.seq:*

5: /cgn2.6/ptodata/2/ina/PTUS_COMB.seq:*

6: /cgn2.6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	604.4	39.1	5865	US-09-011-745-8	Sequence 8, Appli
2	604.4	39.1	6028	US-09-011-745-5	Sequence 5, Appli
3	604.4	39.1	6061	US-09-011-745-6	Sequence 6, Appli
4	604.4	39.1	6312	US-09-011-745-7	Sequence 7, Appli
5	604.2	39.1	13254	US-08-276-852-156	Sequence 156, App
6	604.2	39.1	13254	US-08-276-852-170	Sequence 170, App
7	604.2	39.1	13254	US-08-899-575-156	Sequence 156, App
8	604.2	39.1	13254	US-08-899-575-170	Sequence 170, App
9	604.2	39.1	13254	US-08-899-575-156	Sequence 156, App
10	604.2	39.1	13254	US-08-899-575-170	Sequence 170, App
11	604.2	39.1	13254	PCT-US95-08743-156	Sequence 156, App
12	604.2	39.1	13254	PCT-US95-08743-170	Sequence 170, App
13	603	39.0	1305	US-08-594-469-9	Sequence 9, Appli
14	603	39.0	1905	US-08-906-957-9	Sequence 9, Appli
15	603	39.0	2927	US-08-941-647A-1	Sequence 1, Appli
16	603	39.0	2939	US-08-119-512-2	Sequence 2, Appli
17	603	39.0	2939	US-08-488-015B-2	Sequence 2, Appli
18	603	39.0	2939	US-08-814-412-17	Sequence 17, Appli
19	603	39.0	3003	5182260-18	Patent No. 5182260
20	603	39.0	3104	US-07-415-307A-1	Sequence 1, Appli
21	603	39.0	3104	US-08-371-320-1	Sequence 1, Appli
22	603	39.0	3130	US-09-038-141-1	Sequence 1, Appli
23	603	39.0	3301	US-08-447-430A-42	Sequence 42, Appli
24	603	39.0	3423	US-08-447-430A-40	Sequence 40, Appli
25	603	39.0	3474	US-08-447-430A-41	Sequence 41, Appli
26	603	39.0	3474	US-08-318-837-10	Sequence 10, Appli
27	603	39.0	3585	US-08-362-670B-9	Sequence 9, Appli

C 28	603	39.0	3585	3	US-08-333-576C-9	Sequence 9, Appli
C 29	603	39.0	3585	4	US-08-808-324-9	Sequence 9, Appli
C 30	603	39.0	3585	5	PCT-US94-14030A-9	Sequence 9, Appli
C 31	603	39.0	3623	2	US-07-989-847-13	Sequence 13, Appli
C 32	603	39.0	3623	4	US-08-469-411-13	Sequence 13, Appli
C 33	603	39.0	3632	1	US-07-745-382-13	Sequence 13, Appli
C 34	603	39.0	3632	1	US-07-921-848-13	Sequence 13, Appli
C 35	603	39.0	3632	1	US-08-115-680-3	Sequence 3, Appli
C 36	603	39.0	3632	1	US-07-941-372-3	Sequence 3, Appli
C 37	603	39.0	3632	1	US-08-165-301A-13	Sequence 13, Appli
C 38	603	39.0	3632	3	US-08-810-436-13	Sequence 13, Appli
C 39	603	39.0	3632	5	PCT-US93-08247-3	Sequence 3, Appli
C 40	603	39.0	3632	5	PCT-US94-14179-13	Sequence 13, Appli
C 41	603	39.0	3656	1	US-08-232-463-8	Sequence 8, Appli
C 42	603	39.0	3688	1	US-08-232-463-9	Sequence 9, Appli
C 43	603	39.0	3789	4	US-09-075-019-8	Sequence 8, Appli
C 44	603	39.0	3803	1	US-07-623-953-1	Sequence 1, Appli
C 45	603	39.0	3822	3	US-08-675-566-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-011-745-8
; Sequence 8, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: Construct
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3611)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3612)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3613)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3614)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3799)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3800)
; OTHER INFORMATION: n is any nucleotide

```
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (3801)  
; OTHER INFORMATION: n is any nucleotide  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (3802)  
; OTHER INFORMATION: n is any nucleotide  
US-09-011-745-8
```

Query Match 39.1%; Score 604.4; DB 4; Length 5865;
Best Local Similarity 91.4%; Pred. No. 6.2e-169;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY	595	TTTACTGGCCTGCTCCCTTTATCGGAAGCGGGGCGCATCATATAAATGACGCGCGCGTG	654
DB	3752	ttatcatgtctggattccagatctgaggccaatgagcgcgagatcgaatnnnacatgtgag	3811
QY	655	TAAAGTGTTCAGTTTGAGAAAATTCCTGACGCCGCGCGTGTCTGGCGGTTCCTCAATA	714
DB	3812	cataaaggccagcaaaaggccgaacccgtaaaaaggcgcttgctggcgcttttcccata	3871
QY	715	GGCTCCGCCCCCCTGCACGAGCATCACAAAAATGACAGCTCAAGTCAGAGGTGGCGAAACC	774
DB	3872	ggctccgcccccctgacgagcatcacaaaaatgcactcaagtcaagtgagggaacc	3931
QY	775	CGACAGGACTATAAAGATACACAGCGGTTCCCTCGTGAAGTCCCTCTGTCGGCTCTCTGT	834
DB	3932	cyacagagactataaagataccagcggtttcccccttggaaagtcacctctgcgtctcctg	3991
QY	835	TTCCGACCCCTGCCGCTTACCGGATACCTGTCCGCGCTTTCCTTCGGGAAGCGTGGCGC	894
DB	3992	tccgacctcgcttacogataacctgtccgcctttctcccttcgggaagcgctggcgcc	4051
QY	895	TTTCTCAATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTGCTCCCAAGCTGG	954
DB	4052	tttctcaatgctcacgctgtaggtatctcagttcggttagtgtcgctccaagctgg	4111
QY	955	GCTGTGTGCAGAACCCCCCGTTACGCCGACCGCTGCGCCCTTATCCGGTAACCTATGTC	1014
DB	4112	gctgtgtgcagaaccccccggttcagccgacctgcgcttatccggtaacctatgctc	4171
QY	1015	TTGAGTCAAACCCCGGTGAAGACACGACTATTGCGCACATGGCAGCAGCCACTGGTAACAGA	1074
DB	4172	ttgagtcacaacccoggtaagacacgaacttatcgccaatggcagcgaactgtgaacagg	4231
QY	1075	TTAGCAGAGCGAGTATGTAGCGCGTCTACAGAGTTCCTGGAAGTGTGGCGCTTAACCTACG	1134
DB	4232	ttagcagagcgaggtatgttagcggtgtctacagagttcttgaadtgtgtgctgaactacg	4291
QY	1135	GCATACATAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTCACCTTCGGAA	1194
DB	4292	gctacactagaagacagatttggatatcgctctgctgaagccaagttacctcttcggaa	4351
QY	1195	AAGAGTTGTAGCTCTTGATTCGGGAAACAACACCGCTGTGGTAGCGGTGGTTTTTTTG	1254
DB	4352	aaagagttgtagctcttgccttcggttcggcaaacaccccgctggtagcggtggtttttg	4411
QY	1255	TTTGCAGCAGCAGATATTCGCGCAGAAAAAAGGATCT	1292
DB	4412	tttgcagcagcagatttacgcgcagaaaaaaggatct	4449

RESULT 2
US-09-011-745-5
; Sequence 5, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic

```

> TITLE OF INVENTION: Expression systems
> FILE REFERENCE: 09/011,745
> CURRENT APPLICATION NUMBER: US/09/011,745
> CURRENT FILING DATE: 1998-06-22
> EARLIER APPLICATION NUMBER: PCT/GB96/02061
> EARLIER FILING DATE: 1996-08-23
> EARLIER APPLICATION NUMBER: GB9517263.1
> EARLIER FILING DATE: 1995-08-23
> NUMBER OF SEQ ID NOS: 29
> SOFTWARE: PatentIn Ver. 2.0
> SEQ ID NO 5
> LENGTH: 6028
> TYPE: DNA
> ORGANISM: Artificial Sequence
> FEATURE:
> OTHER INFORMATION: Description of Artificial Sequence: Portion of
> OTHER INFORMATION: construct
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3774)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3775)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3776)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3777)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3962)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3963)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3964)
> OTHER INFORMATION: n is any nucleotide
> FEATURE:
> NAME/KEY: misc.feature
> LOCATION: (3965)
> OTHER INFORMATION: n is any nucleotide
> IS-09-011-745-5

```

Query Match	39.1%	Score	604.4	DB 4	Length	6028
Best Local Similarity	91.4%	Pred. No.	6.3e-169			
Matches	638	Conservative	0	Mismatches	169	
				Indels	0	Gaps
QY	595	TTTACTGGCTGCTCCCTTATTCGGGAAGCGGGCGCATCATATCAATGACGCGCCGCTG	654			
Db	3915	ttatcatgtctggatccagatcttggcccatcgccgcggatcgatnnnacatgtgag	3974			
QY	655	TAAAGTGTTCAGTTGTGAAGAAGAAATTCGTGCAGCCGCGCGTGTCTGGCGTTCCTCCATA	714			
Db	3975	caaaagccagcaaaagccaggaacgcgtaaaaagccggttgcgtttttccata	4034			
QY	715	GGCTCCGCCCCCTGACGAGCATCACAAAATCAGACGCTCAAGTCAGAGTNGCGGAACC	774			
Db	4035	ggtcgcgccccctgacgagcatcaaaaaatgcagcgtcaagtccagaggtggcgcaacc	4094			
QY	775	CGACAGGACTATTAAGATACGAGCTTCCTCCCTGGAGCTCCCTCGTGGCGTCTCCTG	834			
Db	4095	cgacaggaactataaagataccagcgtttccccctgaaagctccctcgtgcgtctcgtg	4154			
QY	835	TTCCGACCCCTCGCGCTTACCGGATACCTGTCCGCGCTTTCCTTCCTCGGGAAGCGTGGCG	894			

RESULT 2
US-09-011-745-5
; Sequence 5, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic

```

; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3995)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3996)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3997)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3998)
; OTHER INFORMATION: n is any nucleotide
; US-09-011-745-6

Query Match      39.1%; Score 604.4; DB 4; Length 6061;
Best Local Similarity 91.4%; Pred. No. 6.3e-169;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAGAGCGGGCGGCATCATATCAAAATGACGCGCGCTG 654
Db 3948 ttatcatgtctggatccagatcttggcccatcgccgcgagatcnnnacaatgtgag 4007
QY 655 TAAAGTCTTACGTTGAGAAAATTCCTCGAGCCCGCGGCTTCTGCTGGCTTTTTCATA 714
Db 4008 caaaagccagcaaaagccaggaacccgtataaaagccggttctgtgctgttttccata 4067
QY 715 GGCTCCGCGCCCTGAGCAGCATCACAAAATCGACCTCAAGTCAGAGTGGCGGAACC 774
Db 4068 ggtctccgccccctgagcagcatcacaaaaatcgacctcaagtcagaggtggcgaaacc 4127
QY 775 CGCAGGACTATAAAGATACACAGCGCTTCCCGCTGGAAGCTCCCTGCTGGCTCTCTG 834
Db 4128 cgacagactataaagataaccagcggttcccccctggaagctccctcgtcgctcctg 4187
QY 835 TTCGACCCCTGCGCTTACCGGATACCTGCGCGCTTTCCTCCCTTCGGGAAGCGTGGCG 894
Db 4188 ttcgacccctgcgcttaccggatcacgtctcgccttctcccttcggaagcggtggcg 4247
QY 895 TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGGTAGTGTCTGCTTCCCAAGCTGG 954
Db 4248 ttctcaatgctcacgctgtaggatctctcagttcgggttctcgtcccaagctgg 4307
QY 955 GCTGTGTCAGCAACCCCGCTTTCAGCCCGACCGCTGCGCTTATCCGGTAACATATGCTC 1014
Db 4308 gctgtgtcagcaaccccccgcttcagcccgacgctgccccttatccgtaactatgctc 4367
QY 1015 TTGAGTCCAAACCCGGTAAGACACACTTATGCGCACCTGGCAGCAGCCACTGGTAACAGGA 1074
Db 4368 ttgagtcacaacccggtaagacacacttatcgccacttggcagcagccactgtgttaacagga 4427
QY 1075 TTACGACGCGAGCTATGTAGCGGCTGCTACAGAGTTCCTGAAGTGTGGCTAACTACG 1134
Db 4428 ttacgagcgaggtatgtaggcggtgctacagagttcttgaagtgggtggcctaactacg 4487
QY 1135 GCTACACTAGAAAGGACAGTATTTGGTATCTGCGCTGCTGCTGAAGCCAGTACCTTCGGA 1194
Db 4488 gctacactagaaggaagatttggatctgctgtcgtcgtggaagccaggttaccttcgaa 4547
QY 1195 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACACCGCTGGTAGCGGTGGTTTTTTT 1254
Db 4548 aaagagttggtagctcttgatcccggaacaaacacccgctggtagcggtggtttttt 4607
QY 1255 TTTGCAAGCAGCAGATTACGCGCGAGAAAAAAGGATCT 1292
Db 4608 ttgcaagcagcagattacgcgcagaaaaaaggatct 4645

```

RESULT 3

```

US-09-011-745-6
; Sequence 6, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Ioic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 6061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3807)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3808)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3809)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3810)

```

```

RESULT      4
US-09-011-745-7
US-09-011-745-7, Application US/09011745
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: Collins, Mary XL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Loic
TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 7
LENGTH: 6312
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
OTHER INFORMATION: construct
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4058)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4059)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4060)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4061)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4246)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4247)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4248)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4249)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-7

```

	Query Match	39.1%	Score 604.4;	DB 4;	Length 6312;
	Best Local Similarity	91.4%;	Pred. No. 6.4e-169;		
	Matches 638;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;
Qy	595 TTTACTGGCGCTGCTCCCTTATTCGGAGACGGGGCGCATCATATCAAAATGACGCCGCCTG	654			
Db	4199 ttatcatgctgtgccagatcgtggcccatgcgcgcgatcgatnnnaccatgtag	4258			
Qy	655 TAAAGTGTTACGTGTGAGAAAGAATTCTGCAGCCCGCGCGTGTGCTGGCGTTTTTTCATTA	714			
Db	4259 caaaaggccagcaaaagccaggacogtaaaaagggccggcttgcgtagggctttttccata	4318			

QY	715	GGCTCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTTCAGAGGTGCGGAAACC	774
Db	4319	ggctcgccccctgacgagcatcaaaaaatcgacgctcaagtcaaggtgcgaaacc	4378
QY	775	CGACGAGCTATAAAGATACCAAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTG	834
Db	4379	cgacgagactaaagataccagcggtttcccccggaaagtcacctgctcctctg	4438
QY	835	TTCCGACCCCTGCGGCTTACCGATACCTGTTCGGCCTTTCCTCCCTTCGGGAAGCGTGGCGC	894
Db	4439	ttccgacctgcccgttaacgataccctgtccgcttctcccttcggaagcgtggtg	4498
QY	895	TTTCTCAATGCTCAGCGCTGTAGGTATCTCAGTTTCGGTGTAGGTGGTTCGCTCCAAAGCTGG	954
Db	4499	tttctcaatgctcacgctgtaggtatctcagttcggtagtgcttgcctccaagctgg	4558
QY	955	GCTGTGTGCACGAACCCCGTTACCGCGGACCGCTGCGCCTTATTCGGGTAACTATCGTC	1014
Db	4559	gctgtgtgcaggaacccccggttcagccgcagcgtgcgcttatccggttaactatcgtc	4618
QY	1015	TTGAGTCCAAACCGGTAAACACAGCACTTATCCCACTGCGCAGCAGCACATGGTAACAGGA	1074
Db	4619	ttgagtcgaacccggttaagacacgactatcgccactggcagcagcactggttaacagga	4678
QY	1075	TTACGACGCGAGGTATGTAGGCGGTGTCTACAGAGTTCTTGAAGTGGTGGCCTTAACTAGC	1134
Db	4679	ttagcagagcgagtgatgtaggcgtgtctacagattcttgaagtgggtgacctaacctacg	4738
QY	1135	GCTACTACTAGAAGACAGATTTTGGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAA	1194
Db	4739	gctactactagaagagacagtatttgggtatctgcgctctgctgaagcaggttaacctcgaa	4798
QY	1195	AAAGAGTTGTTAGCTCTTGATCCGCAAAACAAACACACCGCTGTTAGCGGTGTTTTTTTG	1254
Db	4799	aaagagttggtagctctctgctgcgcgcaaacacacacacacacgctgtgtagcgtgtgtt	4858
QY	1255	TTTGCAGCAGCAGATTACCGCGCAGAAAAAAGGATCT	1292
Db	4859	tttgcagcagcagattaacgcgagaaaaaagatct	4896
RESULT 5			
US-08-276-852-156			
; Sequence 156, Application US/08276852			
; Patent No. 5652138			
; GENERAL INFORMATION:			
; APPLICANT: Burton, Dennis R			
; APPLICANT: Barbas, Carlos F			
; APPLICANT: Lerner, Richard A			
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES			
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS			
; NUMBER OF SEQUENCES: 170			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: The Scripps Research Institute, Office of			
; ADDRESSEE: Patent Counsel			
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,			
; STREET: Mail Drop TPC8			
; CITY: La Jolla			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 92037			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/276.852			
; FILING DATE: 18-JUL-1994			
; CLASSIFICATION: 514			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/178,302			
; FILING DATE: 30-SEP-1993			


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-156

Query Match          39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGCAGCCCGCGCTGCTGGCGTTCCTCATAGGCTCCGCCCC 727
DB 5976 TAAGGGAGAGCGTGCACCTCGGGCGCGTTCCTCATAGGCTCCGCCCC 6035

QY 728 TGAGGACATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGCAGAGACTATA 787
DB 6036 TGAGGACATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGCAGAGACTATA 6095

QY 788 AAGATACAGCGGTTCCCTCGAAGCTCCCTGCTGCTCTCTCTCGACCCGTC 847
DB 6096 AAGATACAGCGGTTCCCTCGAAGCTCCCTGCTGCTCTCTCTCGACCCGTC 6155

QY 848 GCTTACCGGATACCTGTCGCGCTTCCTCCCTCGGGAAGGTGGCGTTCCTCAATGCTC 907
DB 6156 GCTTACCGGATACCTGTCGCGCTTCCTCCCTCGGGAAGGTGGCGTTCCTCAATGCTC 6215

QY 908 ACGCTGTAGGTATCTAGTTCGGTGTAGTTCGCTCCAAAGTGGCGTGTGTGCACGA 967
DB 6216 ACGCTGTAGGTATCTAGTTCGGTGTAGTTCGCTCCAAAGTGGCGTGTGTGCACGA 6275

QY 968 ACCCCCGTTACAGCCGACCGCTGCTTATCCGGTAACTATGCTCTGAGTCCAAACC 1027
DB 6276 ACCCCCGTTACAGCCGACCGCTGCTTATCCGGTAACTATGCTCTGAGTCCAAACC 6335

QY 1028 GGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 1087
DB 6336 GGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 6395

QY 1088 GTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCGCTTAACCTACGCGCTACACTAGAAG 1147
DB 6396 GTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCGCTTAACCTACGCGCTACACTAGAAG 6455

QY 1148 GACAGTATTGGTATCTGCGCTGCTGTAAGCCAGTTCCTCGGAAAGAGTTGTAG 1207
DB 6456 GACAGTATTGGTATCTGCGCTGCTGTAAGCCAGTTCCTCGGAAAGAGTTGTAG 6515

QY 1208 CTCCTGTATCGGCAAAACCAACCGCTGCTGAGGGTGGTTTTTTTGTTCGAAGACGACGA 1267
DB 6516 CTCCTGTATCGGCAAAACCAACCGCTGCTGAGGGTGGTTTTTTTGTTCGAAGACGACGA 6575

QY 1268 GATTACCGCGCAAAAAAAGATCT 1292
DB 6576 GATTACCGCGCAAAAAAAGATCT 6600

RESULT 6
US-08-276-852-170/c
; Sequence 170, Application US/08276852
; Patent No. 5652138

```

```

; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-170

```

```

Query Match          39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGCAGCCCGCGTTCCTCATAGGCTCCGCCCC 727
DB 7279 TAAGGGAGAGCGTGCACCTCGGGCGCGTTCCTCATAGGCTCCGCCCC 7220

QY 728 TGAGGACATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGCAGAGACTATA 787
DB 7219 TGAGGACATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGCAGAGACTATA 7160

QY 788 AAGATACAGCGGTTCCCTCGAAGCTCCCTGCTGCTCTCTCTCGACCCGTC 847
DB 7159 AAGATACAGCGGTTCCCTCGAAGCTCCCTGCTGCTCTCTCTCGACCCGTC 7100

QY 848 GCTTACCGGATACCTGTCGCGCTTCCTCCCTCGGGAAGGTGGCGTTCCTCAATGCTC 907
DB 7099 GCTTACCGGATACCTGTCGCGCTTCCTCCCTCGGGAAGGTGGCGTTCCTCAATGCTC 7040

QY 908 ACGCTGTAGGTATCTAGTTCGGTGTAGTTCGCTCCAAAGTGGCGTGTGTGCACGA 967
DB 7039 ACGCTGTAGGTATCTAGTTCGGTGTAGTTCGCTCCAAAGTGGCGTGTGTGCACGA 6980

QY 968 ACCCCCGTTACAGCCGACCGCTGCTTATCCGGTAACTATGCTCTGAGTCCAAACC 1027

```

Db 6979 ACCCCCGTTACGGCGAGCGCTGCGCCTTATCCGTAATATCGTTCGTAGTCCAAACC 6920
QY 1028 GGTAAAGACAGCACTTATCGCACTTGGCAGCAGCACTGTTAACAGGATTAGCAGCGAG 1087
Db 6919 GGTAAAGACAGCACTTATCGCACTTGGCAGCAGCACTGTTAACAGGATTAGCAGCGAG 6860
QY 1088 GTATGAGCGGTGCTACAGAGTCTTGAAGTGTGGCCTTAACCTAGCGCTACACTAGAAG 1147
Db 6859 GTATGAGCGGTGCTACAGAGTCTTGAAGTGTGGCCTTAACCTAGCGCTACACTAGAAG 6800
QY 1148 GACAGTATTGGTATCTGGCTCTGCTGAAGCCAGTACCTTCGGAAAAAGAGTTGGTAG 1207
Db 6799 GACAGTATTGGTATCTGGCTCTGCTGAAGCCAGTACCTTCGGAAAAAGAGTTGGTAG 6740
QY 1208 CTCTTGATCGGCAAAACCAACCCCTGCTGAGCGTGGTGTGTTTTTTTGTGCAAGCAGCA 1267
Db 6739 CTCTTGATCGGCAAAACCAACCCCTGCTGAGCCAGTACCTTCGGAAAAAGAGTTGGTAG 6680
QY 1268 GATTACGCGCAGAAAAAGGATCT 1292
Db 6679 GATTACGCGCAGAAAAAGGATCT 6655

RESULT 7

US-08-899-575-156
; Sequence 156, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-156
Query Match 39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 668 TGAGAAAGATTCCTCAGCCGCGCGCTTGTGGGGTTTTTCCATAGGCTCGCCGCCCC 727
Db 5976 TAAGGAGAGCGTCGACCTCGGGCCGCGTGTCTGGCGTTTTTCCATAGGCTCGCCGCCCC 6035
QY 728 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 787
Db 6036 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 6095
QY 788 AGATACACAGGGGTTTCCCGCTGGAAAGCTCCCTGCGCTCTCTGTTCCGACCTGCC 847
Db 6096 AAGATACACAGGGGTTTCCCGCTGGAAAGCTCCCTGCGCTCTCTGTTCCGACCTGCC 6155
QY 848 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
Db 6156 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 6215
QY 908 ACGCTGTAGTATCTCAGTTCGGTGTAGTTCGTTCTCCCAAGCTGGGCTGTGTGCACGA 967
Db 6216 ACGCTGTAGTATCTCAGTTCGGTGTAGTTCGTTCTCCCAAGCTGGGCTGTGTGCACGA 6275
QY 968 ACCCGCGTTTCAGCCGCGCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAACC 1027
Db 6276 ACCCGCGTTTCAGCCGCGCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAACC 6335
QY 1028 GGTAAACACAGCACTTATCGCCACTGGCAGCAGCACTGTTAACAGGATTAGCAGCGAG 1087
Db 6336 GGTAAACACAGCACTTATCGCCACTGGCAGCAGCACTGTTAACAGGATTAGCAGCGAG 6395
QY 1088 GTATGAGCGGTGTACAGAGTCTTGAAGTGGTGGCGCTAACTACGGCTACACTAGAAG 1147
Db 6396 GTATGAGCGGTGTACAGAGTCTTGAAGTGGTGGCGCTAACTACGGCTACACTAGAAG 6455
QY 1148 GACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAAAAGAGTTGGTAG 1207
Db 6456 GACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAAAAGAGTTGGTAG 6515
QY 1208 CTCTTGATCGGCAAAACCAACCCCTGCTAGCGGTGGTTTTTTTGTGCAAGCAGCA 1267
Db 6516 CTCTTGATCGGCAAAACCAACCCCTGCTAGCGGTGGTTTTTTTGTGCAAGCAGCA 6575
QY 1268 GATTACGCGCAGAAAAAGGATCT 1292
Db 6576 GATTACGCGCAGAAAAAGGATCT 6600

RESULT 8

US-08-899-575-170/c
; Sequence 170, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA

```

; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-899-575-170

Query Match          39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. le-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0

QY 668 TGAGAAAGAATTCTCGGAGCCCGGGTTCGTGGCGTTTTCATAGGCTCCGCCCCCC 727
DB 7279 TAAGGGAGAGCGTCGACTCGGGCGGGTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC 7220

QY 728 TGACGAGCATCACAAAATCAGCCTCAAGTCAGAGGTGGCGAAACCAGCAGGACTATA 787
DB 7219 TGACGAGCATCACAAAATCAGCCTCAAGTCAGAGGTGGCGAAACCAGCAGGACTATA 7160

QY 788 AAGATACAGCGGTTTCCCCTCGGAAGCTCCCTCGTGGGCTCTCTGTTCGACCCCTGCC 847
DB 7159 AAGATACAGCGGTTTCCCCCTGGAAGCTCCCTCGTGGGCTCTCTGTTCGACCCCTGCC 7100

QY 848 GCTTACCGGATACCTGTCGGCTTTCTCCCTTCGGGAGCGGTGGCGCTTCTCAATGCCTC 907
DB 7099 GCTTACCGGATACCTGTCGGCTTTCTCCCTTCGGGAGCGGTGGCGCTTCTCAATGCCTC 7040

QY 908 ACGCTGATAGGTATCTCAGTTCCGTTAGTTCGCTGCCAAGCTGGGCTGTGTGCACGA 967
DB 7039 ACGCTGATAGGTATCTCAGTTCCGTTAGTTCGCTGCCAAGCTGGGCTGTGTGCACGA 6980

QY 968 ACCCCCGGTTTCAGCCCGACCGCTCGGCTTAATCCGGTAACATATCGTCTTAGTCCAACCC 1027
DB 6979 ACCCCCGGTTTCAGCCCGACCGCTCGGCTTAATCCGGTAACATATCGTCTTAGTCCAACCC 6920

QY 1028 GGTAAGACACGACATTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTACGACAGCGAG 1087
DB 6919 GGTAAGACACGACATTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTACGACAGCGAG 6860

QY 1088 GTATGTAGGCGGTGCTACAGATTCTTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAG 1147
DB 6859 GTATGTAGGCGGTGCTACAGATTCTTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAG 6800

QY 1148 GACAGTATTGGTATATCGCGCTCTGTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 1207
DB 1148 GACAGTATTGGTATATCGCGCTCTGTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 1207

```

QY 728 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATA 787
Db 6036 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATA 6095
QY 788 AAGATACCGAGCTTCCCGCTGGAAGCTCCCTGCGGCTCTCTCTGTTCCGACCCCTGCC 847
Db 6096 AAGATACCGAGCTTCCCGCTGGAAGCTCCCTGCGGCTCTCTCTGTTCCGACCCCTGCC 6155
QY 848 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCCGGAAGGCTGGCGCTTCTCAATGCTC 907
Db 6156 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCCGGAAGGCTGGCGCTTCTCAATGCTC 6215
QY 908 ACCTGTAGTATCTCAGTTCCGCTGAGTGTCTGCTCCAAAGCTGGGCTGTGTGACAGCA 967
Db 6216 ACCTGTAGTATCTCAGTTCCGCTGAGTGTCTGCTCCAAAGCTGGGCTGTGTGACAGCA 6275
QY 968 ACCCCCGTTACCGCGAGCGCTGCGCCTTATCCGCTAACTATCTGCTTCTGAGTCCAAACC 1027
Db 6276 ACCCCCGTTACCGCGAGCGCTGCGCCTTATCCGCTAACTATCTGCTTCTGAGTCCAAACC 6335
QY 1028 GGTAAAGACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 1087
Db 6336 GGTAAAGACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 6395
QY 1088 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAGTACAGGCTTACACTAGAAG 1147
Db 6396 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAGTACAGGCTTACACTAGAAG 6455
QY 1148 GACAGTATTTGGTATCTGCGCTCTGCTGAGCGGTGGTTTTTTTGTGCAAGCAGCA 1267
Db 6456 GACAGTATTTGGTATCTGCGCTCTGCTGAGCGGTGGTTTTTTTGTGCAAGCAGCA 6515
QY 1208 CTCTTGATCGGCAAAACAAACCACCCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCA 1267
Db 6516 CTCTTGATCGGCAAAACAAACCACCCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCA 6575
QY 1268 GATTACGCGCAGAAAAAAGGATCT 1292
Db 6576 GATTACGCGCAGAAAAAAGGATCT 6600

RESULT 10

US-08-899-575-170/C
; Sequence 170, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSER: Patent Counsel
; STREET: 10566 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 08/276,852
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994

; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGCAGCCGCGCTTGTGGGGTTTTTCCATAGGCTCGGCCCC 727
Db 7279 TAAGGAGAGCGTCGACCTCGGGCGCGTGTGGGGTTTTTCCATAGGCTCGGCCCC 7220
QY 728 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATA 787
Db 7219 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATA 7160
QY 788 AAGATACCGAGCGTTTCCCGCTGGAGCTCCCTGCGGCTCTCTCTGTTCCGACCCCTGCC 847
Db 7159 AAGATACCGAGCGTTTCCCGCTGGAGCTCCCTGCGGCTCTCTCTGTTCCGACCCCTGCC 7100
QY 848 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
Db 7099 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 7040
QY 908 AGCTGTAGTATCTCAGTTCCGCTGAGTGTGCTTCCCTCAAGCTGGGCTGTGTGACAGCA 967
Db 7039 AGCTGTAGTATCTCAGTTCCGCTGAGTGTGCTTCCCTCAAGCTGGGCTGTGTGACAGCA 6980
QY 968 ACCCCCGTTACCGCGAGCGCTGGCCCTTATCCGCTAACTATCGTTGAGTCCAAACC 1027
Db 6979 ACCCCCGTTACCGCGAGCGCTGGCCCTTATCCGCTAACTATCGTTGAGTCCAAACC 6920
QY 1028 GGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 1087
Db 6919 GGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 6860
QY 1088 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAGTACAGGCTTACACTAGAAG 1147
Db 6859 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAGTACAGGCTTACACTAGAAG 6800
QY 1148 GACAGTATTTGGTATCTGCGCTCTGCTGAGCGGCTTACCTTCGGAAAAAGAGTTGGTAG 1207
Db 6799 GACAGTATTTGGTATCTGCGCTCTGCTGAGCGGCTTACCTTCGGAAAAAGAGTTGGTAG 6740
QY 1208 CTCTTGATCGGCAAAACAAACCACCCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCA 1267
Db 6739 CTCTTGATCGGCAAAACAAACCACCCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCA 6680
QY 1268 GATTACGCGCAGAAAAAAGGATCT 1292
Db 6679 GATTACGCGCAGAAAAAAGGATCT 6655

RESULT 11

PCT-US95-08743-156

Db 6739 CTCCTGATCCGGCAACAAACACCGCTGGTAGCGTGGTTTTTTTGTTCAGACGACA 6680
Qy 1268 GATTACGGCGCAGAAAAAAGGATCT 1292
Db 6679 GATTACGGCGCAGAAAAAAGGATCT 6655
RESULT 13
US-08-594-469-9
; Sequence 9, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; TITLE OF INVENTION: presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,469
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E.
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-1080
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-594-469-9
Query Match 39.0%; Score 603; DB 1; Length 1905;
Best Local Similarity 100.0%; Pred. No. 9.5e-169;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 690 GCCGCGTTGCTGGCGTTTTCATAGGCTCGGCCCGCTGACGAGCATCAGAAAAATCGA 749
Db 67 GCCGCGTTGCTGGCGTTTTCATAGGCTCGGCCCGCTGACGAGCATCAGAAAAATCGA 126
Qy 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACACGAGCGTTTCCCGCT 809
Db 127 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACACGAGCGTTTCCCGCT 186
Qy 810 GGAAGCTCCCTCGTGGCTCTTCCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 869
Db 187 GGAAGCTCCCTCGTGGCTCTTCCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 246
Qy 870 TTCTCCCTTCGGGAGCGTGGCGCTTCTCAATGCTCACCGCTAGGTATCTCAGTTGC 929

Db 247 TTTCTCCCTTCGGGAACGCTGGCGCTTCTCAATGCTCAGCGCTAGGTATCTCAGTTGC 306
Qy 930 GTGTAGTGTGTTTCGCTCCAAAGCTGGGCTGTGTGCACGAAACCCCGCTTCAGCCGCGCCG 989
Db 307 GTGTAGTGTGTTTCGCTCCAAAGCTGGGCTGTGTGCACGAAACCCCGCTTCAGCCGCGCCG 366
Qy 990 TGGCCCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAACACAGCTTATCCGCA 1049
Db 367 TGGCCCTTATCCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAACACAGCTTATCCGCA 426
Qy 1050 CTGGCAGCAGCCACTGTTAACAGGATTAGCAGAGGATGTAGGCGGTGCTTACAGAG 1109
Db 427 CTGGCAGCAGCCACTGTTAACAGGATTAGCAGAGGATGTAGGCGGTGCTTACAGAG 486
Qy 1110 TTCTGAAAGTGTGGCTTAACCTAGCGCTTACACTAGAGGACAGTATTTGGTATCTGGCT 1169
Db 487 TTCTGAAAGTGTGGCTTAACCTAGCGCTTACACTAGAGGACAGTATTTGGTATCTGGCT 546
Qy 1170 CTGCTGAAAGCCAGTTACCTTCGGAAGAAAGAGTGGTAGCTCTTGATCGGCAAAACACC 1229
Db 547 CTGCTGAAAGCCAGTTACCTTCGGAAGAAAGAGTGGTAGCTCTTGATCGGCAAAACACC 606
Qy 1230 ACCGCTGTAGCGGTGGTTTTTCTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGGA 1289
Db 607 ACCGCTGTAGCGGTGGTTTTTCTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGGA 666
Qy 1290 TCT 1292
Db 667 TCT 669
RESULT 14
US-08-906-957-9
; Sequence 9, Application US/08906957
; Patent No. 5856142
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; TITLE OF INVENTION: presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,957
; FILING DATE: 06-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/594,469
; FILING DATE:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E.
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-1080
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-906-957-9

Query Match          39.0%; Score 603; DB 2; Length 1905;
Best Local Similarity 100.0%; Pred. No. 9.5e-169;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 749
DB 67 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 126

QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 809
DB 127 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 186

QY 810 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCC 869
DB 187 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCC 246

QY 870 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTGC 929
DB 247 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTGC 306

QY 930 GTGTAGTGTGCTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGC 989
DB 307 GTGTAGTGTGCTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGC 366

QY 990 TGGCGCTTATCCGCTAACTACGCTTGTGCTCAACCCGCTTAAGACACGACTTATCGCCA 1049
DB 367 TGGCGCTTATCCGCTAACTACGCTTGTGCTCAACCCGCTTAAGACACGACTTATCGCCA 426

QY 1050 CTGCGACGACGCTGCTTAACAGGATTAAGAGGAGGTATGAGCGGTGTCTACAGAG 1109
DB 427 CTGCGACGACGCTGCTTAACAGGATTAAGAGGAGGTATGAGCGGTGTCTACAGAG 486

QY 1110 TTTCTGAAGTGGTGGCTTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169
DB 487 TTTCTGAAGTGGTGGCTTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546

QY 1170 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1229
DB 547 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 606

QY 1230 ACCGCTGCTAGCGGTGCTTTTGTGTTGCAAGCAGCAGATTAAGCGCAAGAAAAGGA 1289
DB 607 ACCGCTGCTAGCGGTGCTTTTGTGTTGCAAGCAGCAGATTAAGCGCAAGAAAAGGA 666

QY 1290 TCT 1292
DB 667 TCT 669

RESULT 15
US-08-941-647A-1/c
; Sequence 1, Application US/08941647A
; Patent No. 5952211
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, ATSIO
; APPLICANT: TANAKA, TOSHIO
; APPLICANT: MATSUO, YUSHI
; APPLICANT: TANASE, SUMIO
; APPLICANT: FUNATSU, MASAHICO
; APPLICANT: ETO, AKIRA
; TITLE OF INVENTION: METHOD FOR PRODUCING ACTIVE HUMAN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,647A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/533,717
; FILING DATE: 26-SEP-1995
; APPLICATION NUMBER: JP 6-268119
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7335-001-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2927 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-941-647A-1

Query Match          39.0%; Score 603; DB 2; Length 2927;
Best Local Similarity 100.0%; Pred. No. 1.2e-168;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 749
DB 2763 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 2704

QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 809
DB 2703 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 2644

QY 810 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCC 869
DB 2643 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCC 2584

QY 870 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGC 929
DB 2583 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGC 2524

QY 930 GTGTAGTGTGCTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGCGGACCGC 989
DB 2523 GTGTAGTGTGCTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGCGGACCGC 2464

QY 990 TGGCGCTTATCCGCTAACTACGCTTGTGAGTCCAAACCCGCTAAGACACGACTTATCGCCA 1049
DB 2463 TGGCGCTTATCCGCTAACTACGCTTGTGAGTCCAAACCCGCTAAGACACGACTTATCGCCA 2404

QY 1050 CTGCGACGACGCTGCTTAACAGGATTAAGAGGAGGTATGAGCGGTGTCTACAGAG 1109
DB 2403 CTGCGACGACGCTGCTTAACAGGATTAAGAGGAGGTATGAGCGGTGTCTACAGAG 2344

QY 1110 TTTCTGAAGTGGTGGCTTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169
DB 2343 TTTCTGAAGTGGTGGCTTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2284

```

Fri Jan 18 08:27:44 2002

QY 1170 CTGCTGAGCCAGTTACCTTCGGAAGAGAGTTGTTAGTCTTGTATCGGCAAAACAACC 1229
Db 2283 CTGCTGAGCCAGTTACCTTCGGAAGAGAGTTGTTAGTCTTGTATCGGCAAAACAACC 2224
QY 1230 ACCGCTGTAGCGGTGTTTTTTTCTTCAAGCAGCAGATTACGCGCAGAAAAAAGGA 1289
Db 2223 ACCGCTGTAGCGGTGTTTTTTTCTTCAAGCAGCAGATTACGCGCAGAAAAAAGGA 2164
QY 1290 TCT 1292
Db 2163 TCT 2161

Search completed: January 17, 2002, 11:55:46
Job time: 17473 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 15:13:31 ; Search time 11375.6 Seconds
(without alignments)
2383.471 Million cell updates/sec

Title: US-09-242-202A-27

Perfect score: 1547

Sequence: 1 GGTACCTGCCACCATGGCGC.....CTGCACAGCCTCTCCACACA 1547

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

- 1: /cgn2_6/ptodata/2/pna/US06 COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/US06 COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US080 COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US081 COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq:*
- 16: /cgn2_6/ptodata/2/pna/US092 COMB.seq:*
- 17: /cgn2_6/ptodata/2/pna/US093 COMB.seq:*
- 18: /cgn2_6/ptodata/2/pna/US094 COMB.seq:*
- 19: /cgn2_6/ptodata/2/pna/US095A COMB.seq:*
- 20: /cgn2_6/ptodata/2/pna/US095B COMB.seq:*
- 21: /cgn2_6/ptodata/2/pna/US095C COMB.seq:*
- 22: /cgn2_6/ptodata/2/pna/US095D COMB.seq:*
- 23: /cgn2_6/ptodata/2/pna/US096A COMB.seq:*
- 24: /cgn2_6/ptodata/2/pna/US096B COMB.seq:*
- 25: /cgn2_6/ptodata/2/pna/US096C COMB.seq:*
- 26: /cgn2_6/ptodata/2/pna/US096D COMB.seq:*
- 27: /cgn2_6/ptodata/2/pna/US096E COMB.seq:*
- 28: /cgn2_6/ptodata/2/pna/US097A COMB.seq:*
- 29: /cgn2_6/ptodata/2/pna/US097B COMB.seq:*
- 30: /cgn2_6/ptodata/2/pna/US097C COMB.seq:*
- 31: /cgn2_6/ptodata/2/pna/US098 COMB.seq:*
- 32: /cgn2_6/ptodata/2/pna/US099 COMB.seq:*
- 33: /cgn2_6/ptodata/2/pna/US000 COMB.seq:*
- 34: /cgn2_6/ptodata/2/pna/US001 COMB.seq:*
- 35: /cgn2_6/ptodata/2/pna/US002 COMB.seq:*
- 36: /cgn2_6/ptodata/2/pna/US003 COMB.seq:*
- 37: /cgn2_6/ptodata/2/pna/US004 COMB.seq:*
- 38: /cgn2_6/ptodata/2/pna/US005 COMB.seq:*
- 39: /cgn2_6/ptodata/2/pna/US006 COMB.seq:*
- 40: /cgn2_6/ptodata/2/pna/US007 COMB.seq:*
- 41: /cgn2_6/ptodata/2/pna/US008 COMB.seq:*
- 42: /cgn2_6/ptodata/2/pna/US009 COMB.seq:*
- 43: /cgn2_6/ptodata/2/pna/US010 COMB.seq:*

- 44: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
- 45: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
- 46: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
- 47: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
- 48: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
- 49: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
- 50: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
- 51: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
- 52: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
- 53: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
- 54: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
- 55: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
- 56: /cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
- 57: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
- 58: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
- 59: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
- 60: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
- 61: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
- 62: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
- 63: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
- 64: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
- 65: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1547	100.0	1547	16	US-09-242-202-27	Sequence 27, Appl
2	1547	100.0	1547	16	US-09-242-202A-27	Sequence 27, Appl
3	1310.4	84.7	1807	16	US-09-242-202-28	Sequence 28, Appl
4	1310.4	84.7	1807	16	US-09-242-202A-28	Sequence 28, Appl
5	1086.6	70.2	2308	16	US-09-242-202-29	Sequence 29, Appl
6	1086.6	70.2	2308	16	US-09-242-202A-29	Sequence 29, Appl
7	607.4	39.3	7252	16	US-09-238-356-27	Sequence 27, Appl
8	604.2	39.1	7073	8	US-08-480-120-15	Sequence 15, Appl
9	604.2	39.1	7864	8	US-08-480-120-20	Sequence 20, Appl
10	603	39.0	608	16	US-09-242-202-26	Sequence 26, Appl
11	603	39.0	2077	1	PCT-US99-06742-7	Sequence 7, Appl
12	603	39.0	2462	18	US-09-496-445-5	Sequence 5, Appl
13	603	39.0	2462	44	US-60-118-478-5	Sequence 5, Appl
14	603	39.0	2577	3	US-07-919-535C-26	Sequence 26, Appl
15	603	39.0	2704	24	US-09-623-551-27	Sequence 27, Appl
16	603	39.0	2754	5	US-08-114-691-7	Sequence 7, Appl
17	603	39.0	2927	9	US-08-533-717-1	Sequence 1, Appl
18	603	39.0	2927	9	US-08-533-717A-1	Sequence 1, Appl
19	603	39.0	2927	15	US-09-142-481-14	Sequence 14, Appl
20	603	39.0	2939	1	PCT-US94-10146-2	Sequence 2, Appl
21	603	39.0	2939	8	US-08-488-015-2	Sequence 2, Appl
22	603	39.0	2939	15	US-09-114-835-2	Sequence 2, Appl
23	603	39.0	3175	1	PCT-US99-06704-2	Sequence 2, Appl
24	603	39.0	3175	25	US-09-646-075-2	Sequence 2, Appl
25	603	39.0	3301	8	US-08-447-430-42	Sequence 42, Appl
26	603	39.0	3301	17	US-09-342-673-42	Sequence 42, Appl
27	603	39.0	3327	21	US-09-546-411A-56	Sequence 56, Appl
28	603	39.0	3327	21	US-09-547-939A-56	Sequence 56, Appl
29	603	39.0	3327	24	US-09-628-730-56	Sequence 56, Appl
30	603	39.0	3327	45	US-60-128-830-56	Sequence 56, Appl
31	603	39.0	3331	15	US-09-148-920-1	Sequence 1, Appl
32	603	39.0	3344	18	US-09-422-563-8	Sequence 8, Appl
33	603	39.0	3369	21	US-09-546-411A-48	Sequence 48, Appl
34	603	39.0	3369	21	US-09-547-939A-48	Sequence 48, Appl
35	603	39.0	3369	24	US-09-628-730-48	Sequence 48, Appl
36	603	39.0	3369	45	US-60-128-830-48	Sequence 48, Appl
37	603	39.0	3420	25	US-09-643-192-31	Sequence 31, Appl
38	603	39.0	3421	15	US-09-130-115-31	Sequence 31, Appl
39	603	39.0	3423	8	US-08-447-430-40	Sequence 40, Appl
40	603	39.0	3423	17	US-09-342-673-40	Sequence 40, Appl
41	603	39.0	3423	17	US-09-342-673-40	Sequence 40, Appl

c 42 603 39.0 3423 18 US-09-485-737A-86 Sequence 86, Appl
c 43 603 39.0 3423 18 US-09-485-737B-86 Sequence 86, Appl
44 603 39.0 3440 3 US-07-671-058-7 Sequence 7, Appl
45 603 39.0 3462 29 US-09-742-373-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-242-202-27
; Sequence 27, Application US/09242202
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202
; FILING DATE: 01-Nov-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1547 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-242-202-27

Query Match 100.0%; Score 1547; DB 16; Length 1547;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACATGGCGCGGATCTTTTATCACTGATAAGTTGGTGACATATTATGTT 60
Db 1 GGTACCTGCCACATGGCGCGGATCTTTTATCACTGATAAGTTGGTGACATATTATGTT 60
QY 61 TATCAGTGATAAGTGTCAAGCATGACAAAGTTGACGCGAATACAGTATCGTCCCGG 120
Db 61 TATCAGTGATAAGTGTCAAGCATGACAAAGTTGACGCGAATACAGTATCGTCCCGG 120
QY 121 CCTGTGACTGTTGAACGAGGTGCGCGTAGACGCTCTGACGACACGCAAACTGCGCGAAG 180
Db 121 CCTGTGACTGTTGAACGAGGTGCGCGTAGACGCTCTGACGACACGCAAACTGCGCGAAG 180

QY 181 GTTGGGGTGCAGCAGCGCGCTTTTACTGGCACTTCAGGAACAAGCGGCGCCCTTAAGG 240
Db 181 GTTGGGGTGCAGCAGCGCGCTTTTACTGGCACTTCAGGAACAAGCGGCGCCCTTAAGG 240
QY 241 GCCATATGTTGAGTGGATGCTTACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
Db 241 GCCATATGTTGAGTGGATGCTTACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCGGCGCAGCACAGGCAATGCCGCTCTTCCCTGCAGGATGAGTAGTGCCCTC 360
Db 301 CCCCGGCGCAGCACAGGCAATGCCGCTCTTCCCTGCAGGATGAGTAGTGCCCTC 360
QY 361 TCCTGGCCCTTGAAGTTGCCACTCCAGTGCACACGAGCTTTCCTAATAAATAAATTAAAGTT 420
Db 361 TCCTGGCCCTTGAAGTTGCCACTCCAGTGCACACGAGCTTTCCTAATAAATAAATTAAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGGTGCTCTATATATATAAGCTGATATCAATCTTTT 480
Db 421 GCATCATTTTGTCTGACTAGGTGCTCTATATATATAAGCTGATATCAATCTTTT 480
QY 481 CGGACTTTTGAAGTGATGGTGGTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGCG 540
Db 481 CGGACTTTTGAAGTGATGGTGGTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGCG 540
QY 541 AGATTAGAGTCTGCTCCTTTTGGCGGCTCGGGAACCCACACGCGGTAATGCTTTTACT 600
Db 541 AGATTAGAGTCTGCTCCTTTTGGCGGCTCGGGAACCCACACGCGGTAATGCTTTTACT 600
QY 601 GGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAGT 660
Db 601 GGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAGT 660
QY 661 GTTAGCTTGAGAAAGAAATTCCTGCAGCGCGCGCTTGTGCGCTTTTTCATAGGCTCC 720
Db 661 GTTAGCTTGAGAAAGAAATTCCTGCAGCGCGCGCTTGTGCGCTTTTTCATAGGCTCC 720
QY 721 GCCCGCTGCAGGATCACAAAATCGAGCTCAAGTCAGAGTGGCGCAACCCACAG 780
Db 721 GCCCGCTGCAGGATCACAAAATCGAGCTCAAGTCAGAGTGGCGCAACCCACAG 780
QY 781 GACTATAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTCTGTTCCGA 840
Db 781 GACTATAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTCTGTTCCGA 840
QY 841 CCTGCGCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTC 900
Db 841 CCTGCGCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTC 900
QY 901 AATGCTCAGCTGTAGGTATCTCAGTTCGCTGCTAGTTCGCTCCAAAGCTGGGCTGTG 960
Db 901 AATGCTCAGCTGTAGGTATCTCAGTTCGCTGCTAGTTCGCTCCAAAGCTGGGCTGTG 960
QY 961 TGCACGAACCCCGCTTACGCGCGGCTGCGCTTATCCGCTAACTATCGTCTTTGAGT 1020
Db 961 TGCACGAACCCCGCTTACGCGCGGCTGCGCTTATCCGCTAACTATCGTCTTTGAGT 1020
QY 1021 CCAACCCGCTTAAGACACGACTTATCCGCACTGCGACACGCGCTGCTAAGGATTTAGCA 1080
Db 1021 CCAACCCGCTTAAGACACGACTTATCCGCACTGCGACACGCGCTGCTAAGGATTTAGCA 1080
QY 1081 GAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCGCTTAACCGGTACA 1140
Db 1081 GAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCGCTTAACCGGTACA 1140
QY 1141 CTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAG 1200
Db 1141 CTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAG 1200
QY 1201 TTGGTAGCTCTTGATCCGCAACCAACCGCTGCTGAGCGGTGTTTTTTTGTGTTGCA 1260
Db 1201 TTGGTAGCTCTTGATCCGCAACCAACCGCTGCTGAGCGGTGTTTTTTTGTGTTGCA 1260
QY 1261 AGCAGCAGATTACGCCCGCAGAAAAAAGGATCTGGGGATCCCGGAGAGCTCACTCTAGATG 1320

Db 1261 ACAGACAGATTAGCCGCAAAAAAGGATCTGGGGATCCGGAGAGCTCACTCTAGATG 1320
QY 1321 AGAGACAGTGGAGGAGACAGAGACTCGAATTTCCGGAGCTATTTCAGTTTCTTTTC 1380
Db 1321 AGAGACAGTGGAGGAGACAGAGACTCGAATTTCCGGAGCTATTTCAGTTTCTTTTC 1380
QY 1381 CGTTTTGTCAATTTCACTTATGATACCGGCCAATCTTGGTTGCTATTTTGGAAACTCC 1440
Db 1381 CGTTTTGTCAATTTCACTTATGATACCGGCCAATCTTGGTTGCTATTTTGGAAACTCC 1440
QY 1441 CCTTAGGGATGCCCTCAACTGGCCCTTAAAGAGGCCAGCTGAGCTGAGAGGATTC 1500
Db 1441 CCTTAGGGATGCCCTCAACTGGCCCTTAAAGAGGCCAGCTGAGCTGAGAGGATTC 1500
QY 1501 TGCAGAGGATCAAGACAGACAGCTGGACCTCGCACACGCTCTCCACA 1547
Db 1501 TGCAGAGGATCAAGACAGACAGCTGGACCTCGCACACGCTCTCCACA 1547

RESULT 2
US-09-242-202A-27
; Sequence 27, Application US/09242202A
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242, 202A
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; LENGTH: 1547 base pairs
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-242-202A-27

Query Match 100.0%; Score 1547; DB 16; Length 1547;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTACCTGCCACCATGGCGGGATTCTTTATCACTGATAAGTTGGTGGACATATTATGTT 60

Db 1 GGTACCTGCCACCATGGCGGGATTCTTTATCACTGATAAGTTGGTGGACATATTATGTT 60
QY 61 TATCAGTGATAAAGTCTCAAGCATGACAAAGTTGCAGCCGAATACAGTGATCGTGCCGG 120
Db 61 TATCAGTGATAAAGTCTCAAGCATGACAAAGTTGCAGCCGAATACAGTGATCGTGCCGG 120
QY 121 CCCTGGACTGTTTGAACGAGGTGCGCGTAGACGCTGTGACGACACGCAAACTGGCGAAACG 180
Db 121 CCCTGGACTGTTTGAACGAGGTGCGCGTAGACGCTGTGACGACACGCAAACTGGCGAAACG 180
QY 181 GTTGGGGGTGCACAGCGCGCTTTACTGGCACTTTCAGGAAACAAGCGGCGCTTTAAGG 240
Db 181 GTTGGGGGTGCACAGCGCGCTTTACTGGCACTTTCAGGAAACAAGCGGCGCTTTAAGG 240
QY 241 GCCATATGGTGTAGTGATGCTTGTACCCAGCGGGGATGGGGAGACCTGTAGTCAGAG 300
Db 241 GCCATATGGTGTAGTGATGCTTGTACCCAGCGGGGATGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCCGGGACAGACAGGCCAATGCCGCTCTTCCCTTCGACGATGAGTAGTGAGTGCCCTC 360
Db 301 CCCCCGGGACAGACAGGCCAATGCCGCTCTTCCCTTCGACGATGAGTAGTGAGTGCCCTC 360
QY 361 TCCTGGCCCTGGAAAGTTGCCACTCCAGTGCCCAACGACCTTGTCCCTAATAAATTAAGTT 420
Db 361 TCCTGGCCCTGGAAAGTTGCCACTCCAGTGCCCAACGACCTTGTCCCTAATAAATTAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGTGCTCTCTATAATATATATAGCTTGATATCGAATCTTT 480
Db 421 GCATCATTTTGTCTGACTAGTGCTCTCTATAATATATAGCTTGATATCGAATCTTT 480
QY 481 CGGACTTTTGAAGAGTGTGGTGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
Db 481 CGGACTTTTGAAGAGTGTGGTGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
QY 541 AGATTTAGAGTCTCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
Db 541 AGATTTAGAGTCTCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
QY 601 GGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGGCGCTGTAAGT 660
Db 601 GGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGGCGCTGTAAGT 660
QY 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGGCGTTTTCCTCATAGGCTCC 720
Db 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGGCGTTTTCCTCATAGGCTCC 720
QY 721 GCCCCCTTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG 780
Db 721 GCCCCCTTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG 780
QY 781 GACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTGTTCCGA 840
Db 781 GACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTGTTCCGA 840
QY 841 CCCTGCGCTTACGGGATACCTGTGCGCCCTTCTCCCTCGGGAAGCGTGCGCTTCTC 900
Db 841 CCCTGCGCTTACGGGATACCTGTGCGCCCTTCTCCCTCGGGAAGCGTGCGCTTCTC 900
QY 901 AATGCTCACGCTGTAGGTATCTCAGTTGCGTGTAGGTCGTTCCGCTCAGGCTGTG 960
Db 901 AATGCTCACGCTGTAGGTATCTCAGTTGCGTGTAGGTCGTTCCGCTCAGGCTGTG 960
QY 961 TGCACGAACCCCGCTTACGCGGAGCGCTTATCCGCTAATATCGCTTTCGAGT 1020
Db 961 TGCACGAACCCCGCTTACGCGGAGCGCTTATCCGCTAATATCGCTTTCGAGT 1020
QY 1021 CCAACCGGTAAAGACAGACGACTTATCGCCACTGCGACGACGACTGGTAACAGGATTAGCA 1080
Db 1021 CCAACCGGTAAAGACAGACGACTTATCGCCACTGCGACGACGACTGGTAACAGGATTAGCA 1080
QY 1081 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACAGGCTACA 1140
Db 1081 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACAGGCTACA 1140

QY 961 TGCACGAACCCCGGTTACGCCGACCGCTTATCCCGGTAACACTATCGTCTTGAGT 1020
Db 961 TGCACGAACCCCGGTTACGCCGACCGCTTATCCCGGTAACACTATCGTCTTGAGT 1020
QY 1021 CCAACCCGGTAAGACACGACTTATCCGCTACTGCGACGACCTGTAACAGATTAGCA 1080
Db 1021 CCAACCCGGTAAGACACGACTTATCCGCTACTGCGACGACCTGTAACAGATTAGCA 1080
QY 1081 GAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGGTACA 1140
Db 1081 GAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGGTACA 1140
QY 1141 CTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTCGGAAAAAGAG 1200
Db 1141 CTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTCGGAAAAAGAG 1200
QY 1201 TTGGTAGCTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTTGGTTGCA 1260
Db 1201 TTGGTAGCTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTTGGTTGCA 1260
QY 1261 AGCAGCAGATTAGCGCCAGAAAAAAGGATCTGGGGGATCCGGAGAGCTCAC 1312
Db 1261 AGCAGCAGATTAGCGCCAGAAAAAAGGATCTGGGGGATCCGGAGAGCTCCC 1312

RESULT 4
US-09-242-202A-28
; Sequence 28, Application US/09242202A
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; APPLICANT: Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202A
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELE: 421792
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-242-202A-28

Query Match 84.7%; Score 1310.4; DB 16; Length 1807;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGTACCTGCCACCATCGCGCGGATTCTTTATCACTGATAGTTGGTGGACATATTATGTT 60
Db 1 GGTACCTGCCACCATCGCGCGGATTCTTTATCACTGATAGTTGGTGGACATATTATGTT 60
QY 61 TATCAGTGATAAGTGTCAAGCATGACAAAGTTGCAGCCGGAATACAGTATCGTGCCTCG 120
Db 61 TATCAGTGATAAGTGTCAAGCATGACAAAGTTGCAGCCGGAATACAGTATCGTGCCTCG 120
QY 121 CCTGTGACTGTTGAACGAGGTGCGCGTAGACGCTCTGACGACACGCAAACTGCGGGAACG 180
Db 121 CCTGTGACTGTTGAACGAGGTGCGCGTAGACGCTCTGACGACACGCAAACTGCGGGAACG 180
QY 181 GTTGGGGGTGCAGCAGCCGCGCTTTACTGGCACTTCAGGAACAAGCGGCGCTTTAAGG 240
Db 181 GTTGGGGGTGCAGCAGCCGCGCTTTACTGGCACTTCAGGAACAAGCGGCGCTTTAAGG 240
QY 241 GCCATATGTTGAGTGTGATGCCCTTGACCCGAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
Db 241 GCCATATGTTGAGTGTGATGCCCTTGACCCGAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCCGGGACGACAGGCCCAATGCCGCTCTTCCCTGCAGGATGAGTAGTGTGCTC 360
Db 301 CCCCCGGGACGACAGGCCCAATGCCGCTCTTCCCTGCAGGATGAGTAGTGTGCTC 360
QY 361 TCCTGGCCCTTGGAAAGTTGCCACTCCAGTGCACCGCCAGCTTGTCTTAATAAATAAAGTT 420
Db 361 TCCTGGCCCTTGGAAAGTTGCCACTCCAGTGCACCGCCAGCTTGTCTTAATAAATAAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGTGTCTCTATATATATATATATATATATATATATATATAT 480
Db 421 GCATCATTTTGTCTGACTAGTGTCTCTATATATATATATATATATATATATATATATAT 480
QY 481 CGGACTTTTGAAGAGTATGTTGGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
Db 481 CGGACTTTTGAAGAGTATGTTGGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
QY 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCGGTAATGCTTTTACT 600
Db 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCGGTAATGCTTTTACT 600
QY 601 GGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATACGCGCGCTGTAAGT 660
Db 601 GGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATACGCGCGCTGTAAGT 660
QY 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGCGGTTTTTCCATAGGCTCC 720
Db 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGCGGTTTTTCCATAGGCTCC 720
QY 721 GCGCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAG 780
Db 721 GCGCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAG 780
QY 781 GACTATAAAGATACGAGCGCTTTCCCGCTGGAAGCTCCCTCGCTCTCTCTGTTCCGA 840
Db 781 GACTATAAAGATACGAGCGCTTTCCCGCTGGAAGCTCCCTCGCTCTCTCTGTTCCGA 840
QY 841 CCTGCGCGCTTACGGGATACCTGTGCGGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTC 900
Db 841 CCTGCGCGCTTACGGGATACCTGTGCGGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTC 900
QY 901 AATGCTCACGCTGAGTATCTCAGTTCGGTGTAGTGTCTCGCTCCCAAGCTGGGCTGTG 960
Db 901 AATGCTCACGCTGAGTATCTCAGTTCGGTGTAGTGTCTCGCTCCCAAGCTGGGCTGTG 960
QY 961 TGCACGAACCCCGGTTACGCCGACCGCTGCGCTTATCCGGTAACACTATCGTCTTGAGT 1020
Db 961 TGCACGAACCCCGGTTACGCCGACCGCTGCGCTTATCCGGTAACACTATCGTCTTGAGT 1020

QY	1021	CCAA	CCCGGTAA	GACACG	AGCTTAT	CGCCACT	GTGGCAG	CAGCCAC	TGTGTA	ACAGGATT	AGCA	1080
Db	1021	CCAA	CCCGGTAA	GACACG	AGCTTAT	CGCCACT	GTGGCAG	CAGCCAC	TGTGTA	ACAGGATT	AGCA	1080
QY	1081	GAGC	GAGGTAT	GTAG	CGGTGCT	ACAGAG	TTCCT	GAACTGT	GCCTAA	CTACGGCT	TACA	1140
Db	1081	GAGC	GAGGTAT	GTAG	CGGTGCT	ACAGAG	TTCCT	GAACTGT	GCCTAA	CTACGGCT	TACA	1140
QY	1141	CTAG	AAGGAC	AGTAT	TGTGTAT	CTCGGCT	CTGCTG	GAAGCC	AGTTAC	CTTCGGA	AAAAAG	1200
Db	1141	CTAG	AAGGAC	AGTAT	TGTGTAT	CTCGGCT	CTGCTG	GAAGCC	AGTTAC	CTTCGGA	AAAAAG	1200
QY	1201	TTGG	TAGCTCT	TTGAT	CCGGCA	AAACAA	CCCGCT	TGTTAG	CGGTGG	TTTTTT	TGTTTGCA	1260
Db	1201	TTGG	TAGCTCT	TTGAT	CCGGCA	AAACAA	CCCGCT	TGTTAG	CGGTGG	TTTTTT	TGTTTGCA	1260
QY	1261	AGCA	GAGATT	TAC	GGCAGAAAA	AAAAAG	GATCT	GTGGGGG	ATCCG	GAGAGCT	CAAC	1312
Db	1261	AGCA	GAGATT	TAC	GGCAGAAAA	AAAAAG	GATCT	GTGGGGG	ATCCG	GAGAGCT	CAAC	1312

RESULT 5
 US-09-242-202-29
 ; Sequence 29, Application US/09242202
 ; GENERAL INFORMATION:
 ; APPLICANT: Nelson, Edward L.
 ; Nelson, Peter J.
 ; TITLE OF INVENTION: NOVEL VECTOR FOR
 ; POLYNUCLEOTIDE VACCINES
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 ; STREET: 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: MS WORD 97
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/242.202
 ; FILING DATE: 01-NOV-1999
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US60/023931
 ; FILING DATE: 14-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KATHRYN M. BROWN
 ; REGISTRATION NUMBER: 34556
 ; REFERENCE/DOCKET NUMBER: 2026-42360US1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; TELEX: 421792
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2308 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: CDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: No
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
 US-09-242-202-29

QY	220	GAACAAGCGGGCGCCTTTAAGGGCCATATGTGAGTGGATGCTTGAACCCAGCGCGGGAT	279
Db	721		
QY	280	GGGGAGACCTGTAGTCAGAGCCCCCGGGCAGCACAGGCCAATGCCCCGTCTTCCCTGCG	339
Db	781		
QY	340	AGGATGAGTAGTGAGTGCCCTCTCTGGCCCTGGAAGTTGCCATCTCCAGTGCCCCACCGC	399
Db	841		
QY	400	TTGTCCTAATAAAATTAAAGTTGTCATATTTGTCTGACTAGGTGTCTCTATAAATTAT	459
Db	901		
QY	460	AAGCTTGATATCGAATTCCTTTCCGACCTTTTGAAGTGATGGTGGGGGAAGGATTCGA	519
Db	961		
QY	520	ACCTTCGAAGTCGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCCGTCGGGAACCC	579
Db	1021		
QY	580	ACCACGGGTAAATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCA	639
Db	1081		
QY	640	AATGACGGCCGTGTAAAGTGTTACGTTGAGAAAGAAATTCCTGCAGCCGCCGCGTTGC	699
Db	1141		
QY	700	TGGCGTTTTTCCATAGGCTCCGCCCCCTTCACGAGCATCACAAAATCGAGCTCAAGTC	759
Db	1201		
QY	760	AGAGTGGCGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCTCTGGAAGCTCC	819
Db	1261		
QY	820	TCGTGCGCTCTCCTGTTCCGACCCCTGCCCTTACCGGATACCTGTCCGCTTTCCTCC	879
Db	1321		
QY	880	CGGGAACGTTGGCGCTTCTCAATGCTACGCTGAGGTATCTCAGTTCGGTGTAGTCTG	939
Db	1381		
QY	940	TTCGCTCCAAGCTGGGCTGTGTGCAGAACCCCGTTCAGCCGACCGCTGCGCGCTTAT	999
Db	1441		
QY	1000	CCGGTAACCTATCTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCACTGGCAGCAG	1059
Db	1501		
QY	1060	CCACTGGTAACAGGATTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGT	1119
Db	1561		
QY	1120	GGTGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTCGCTCTGCTGAAGC	1179
Db	1621		
QY	1180	CAGTTACCTTCGAAAAAGATTGGTAGCTTTGATCGGCAAAACCAACCGCTGSTA	1239
Db	1681		
QY	1240	GCAGTGGTTTTTTTGTTCGACAGCAGCATTTACCGCGAGAAAAAAGGATCTCGGGGAT	1299
Db	1741		
QY	1300	CCGAGAGCTCAC	1312

```
Query Match          70.2%; Score 1086.6; DB 16; Length 2308;
Best Local Similarity 99.6%; Pred. No. 9.6e-310;
Matches 1089; Conservative 0; Mismatches 4; Indels 0;
```

D	b	1801	CCGAGAGCTCCC	1813
RESULT 6				
US-09-242-202A-29				
; Sequence 29, Application US/09242202A				
; GENERAL INFORMATION:				
; APPLICANT: Nelson, Edward L.				
; APPLICANT: Nelson, Peter J.				
; TITLE OF INVENTION: NOVEL VECTOR FOR POLYNUCLEOTIDE VACCINES				
; NUMBER OF SEQUENCES: 29				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.				
; STREET: 345 PARK AVENUE				
; CITY: NEW YORK				
; STATE: NEW YORK				
; COUNTRY: USA				
; ZIP: 10154				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB				
; COMPUTER: IBM PC COMPATIBLE				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: MS WORD 97				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/09/242.202A				
; FILING DATE: 20-Apr-2000				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: US60/023931				
; FILING DATE: 14-AUG-1996				
; ATTORNEY/AGENT INFORMATION:				
; NAME: KATHRYN M. BROWN				
; REGISTRATION NUMBER: 34556				
; REFERENCE/DOCKET NUMBER: 2026-4236US1				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: (212) 758-4800				
; TELEFAX: (212) 751-6849				
; TELEX: 421792				
; INFORMATION FOR SEQ ID NO: 29:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 2308 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: single				
; TOPOLOGY: unknown				
; MOLECULE TYPE: cDNA				
; HYPOTHEetical: No				
; ANTI-SENSE: No				
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:				
US-09-242-202A-29				
Query Match 70.2%; Score 1086.6; DB 16; Length 2308;				
Best Local Similarity 99.6%; Pred. No. 9.6e-310;				
Matches 1089; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				
Qy	220	GAAACAGCGGCCTTAAGGCCATATGGTGAGTGGCTTGACCCCAAGCGGGAT	279	
D	721	GTACAAGTAGCGCCTTAAGGCCATATGGTGAGTGGCTTGACCCCAAGCGGGAT	780	
Qy	280	GGGGAGACCTGTAGTCAGAGCCCCCGGGCAGCACAGGCCAATGCCCTCTCCCTGC	339	
D	781	GGGGAGACCTGTAGTCAGAGCCCCCGGGCAGCACAGGCCAATGCCCTCTCCCTGC	840	
Qy	340	AGGATGAGTAGTGAGTGCCTCTCTGGCCCTGGAGTTGCCACTCCAGTCCCAAGCC	399	
D	841	AGGATGAGTAGTGAGTGCCTCTCTGGCCCTGGAGTTGCCACTCCAGTCCCAAGCC	900	
Qy	400	TTGTCCTAATAAAAATTAAAGTGCATCATTTGTCTGACTAGTGTCCTCTAATATAT	459	
D	901	TTGTCCTAATAAAAATTAAAGTGCATCATTTGTCTGACTAGTGTCCTCTAATATAT	960	
Qy	460	AAGCTTGATATCGAATTTCTTCGGACTTTTGAAGATGATGGTGGGGAGGATTCGA	519	

```

RESULT 7
US-09-238-356-27
; Sequence 27, Application US/09238356
; GENERAL INFORMATION:
; APPLICANT: Kingsman, et al
; TITLE OF INVENTION: Retroviral Vectors
; FILE REFERENCE: 674523-2006
; CURRENT APPLICATION NUMBER: US/09/238,356
; CURRENT FILING DATE: 1999-03-27
; PRIOR APPLICATION NUMBER: PCT/CB/03876
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ IDS NOS: 64
; SOFTWARE: PatentIn version 3.0

```


ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 2026-4236US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-242-202-26

Query Match 39.0%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 6e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCTTGGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATACAAAAATCGA 749
Db |
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCT 809
Db |
QY 66 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCT 125
Db |
QY 810 GGAAGCTCCCTGTCGCTCTCTCTCCGACCTGCGCTTACCGGATACCTGTCGCC 869
Db |
QY 126 GGAAGCTCCCTGTCGCTCTCTCTCCGACCTGCGCTTACCGGATACCTGTCGCC 185
QY 870 TTTCTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTGC 929
Db |
QY 186 TTTCTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTGC 245
QY 930 GTGTAGGTCGTTCCGCTCCAAAGTGGCTGTGTCACGAAACCCCGTTACGCCGACCGC 989
Db |
QY 246 GTGTAGGTCGTTCCGCTCCAAAGTGGCTGTGTCACGAAACCCCGTTACGCCGACCGC 305
QY 990 TGCCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTTACGCCGACCGTATCGCCA 1049
Db |
QY 306 TGCCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTTACGCCGACCGTATCGCCA 365
QY 1050 CTGCAGCAGCCACTGGTAACAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG 1109
Db |
QY 366 CTGCAGCAGCCACTGGTAACAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG 425
QY 1110 TTTCTTGAAGTGGTGGCTTAAGTACAGCTACAGTACAGTATTTGGTATCTCGCT 1169
Db |
QY 426 TTTCTTGAAGTGGTGGCTTAAGTACAGCTACAGTACAGTATTTGGTATCTCGCT 485
QY 1170 CTGCTGAAGCAGTACCTTCGGAAGAGTTGGTGTAGCTTGTATCCGCGAAACAAAC 1229
Db |
QY 486 CTGCTGAAGCAGTACCTTCGGAAGAGTTGGTGTAGCTTGTATCCGCGAAACAAAC 545
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGTTCGACGACGAGATTACGCGCAAAAAAGGA 1289
Db |
QY 546 ACCGCTGGTAGCGGTGTTTTTTTGTTCGACGACGAGATTACGCGCAAAAAAGGA 605
QY 1290 TCT 1292
Db |
QY 606 TCT 608

RESULT 11
US-09-242-202a-26
; Sequence 26, Application US/09242202a

GENERAL INFORMATION:
APPLICANT: Nelson, Edward L.
TITLE OF INVENTION: NOVEL VECTOR FOR
POLYNUCLEOTIDE VACCINES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/242, 202A
FILING DATE: 20-Apr-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US60/023931
FILING DATE: 14-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 2026-4236US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-242-202a-26

Query Match 39.0%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 6e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCTTGGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATACAAAAATCGA 749
Db |
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCT 809
Db |
QY 66 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCT 125
QY 810 GGAAGCTCCCTGTCGCTCTCTCTCCGACCTGCGCTTACCGGATACCTGTCGCC 869
Db |
QY 126 GGAAGCTCCCTGTCGCTCTCTCTCCGACCTGCGCTTACCGGATACCTGTCGCC 185
QY 870 TTTCTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTGC 929
Db |
QY 186 TTTCTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTGC 245
QY 930 GTGTAGGTCGTTCCGCTCCAAAGTGGCTGTGTCACGAAACCCCGTTACGCCGACCGC 989
Db |
QY 246 GTGTAGGTCGTTCCGCTCCAAAGTGGCTGTGTCACGAAACCCCGTTACGCCGACCGC 305
QY 990 TGCCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTTACGCCGACCGTATCGCCA 1049
Db |
QY 306 TGCCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCCCGTTACGCCGACCGTATCGCCA 365

QY 1050 CTGCGACGACCACTGCTAAACAGAGTTAGCAGACGAGGATGTAGCGGCTGCTACAGAG 1109
Db 366 CTGCGACGACCACTGCTAAACAGAGTTAGCAGACGAGGATGTAGCGGCTGCTACAGAG 425
QY 1110 TTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGGACAGTATTGGTATCTGCGCT 1169
Db 426 TTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGGACAGTATTGGTATCTGCGCT 485
QY 1170 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACC 1229
Db 486 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACC 545
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGGTGAAGCAGCAGAGTATACGCGCAGAAAAAGGA 1289
Db 546 ACCGCTGGTAGCGGTGTTTTTTTGGTGAAGCAGCAGAGTATACGCGCAGAAAAAGGA 605
QY 1290 TCT 1292
Db 606 TCT 608
RESULT 12
PCT-US99-06742-7
; Sequence 7, Application PC/TUS9906742
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/06742
; CURRENT FILING DATE: 1999-03-28
; EARLIER APPLICATION NUMBER: 60/079,792
; EARLIER FILING DATE: 1998-03-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Word97
; SEQ ID NO 7
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PASLib
PCT-US99-06742-7

Query Match 39.0%; Score 603; DB 1; Length 2077;
Best Local Similarity 100.0%; Pred. No. 1e-166;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCCGCGTTGCTGGCGTTTTCATAGGCTCGCGCCCTCGACGAGCATCACAAAAATCGA 749
Db 1437 gccgctgtgcgctgttttccataggtctcgcccccctgacgagcatcaaaaaatcga 1496
QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACGAGCTATAAAGATACCGAGCGTTTCCCGCT 809
Db 1497 cgctcaagtcagaggtgggaaacccgacagactataaagataccaggcggtttcccccct 1556
QY 810 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCC 869
Db 1557 ggaagctccctcgctgctctctgttccgacctgcttaaccgatacactgtccgcc 1616
QY 870 TTCTTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTAGTAGTATCTCAGTTCG 929
Db 1617 ttcttcccttcgggaagcggtggcgctttctcaatgctcacgctgtagtatactcagttcg 1676
QY 930 GTGAGGTGCTGCTCCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGC 989
Db 1677 gtgaggtgctgctcccaagctggcgctgtgtgcacgaaccccgcttcagcccgacccg 1736
QY 990 TGGCCCTTATCCGCTAACTATCTGCTTGTAGTCCAAACCGGGTAAGACAGACTTATCGCCA 1049
Db 1737 tggcccttataccgctaaactatctgtttagtccaaacccggtaagacacgacttatacgcca 1796

QY 1050 CTGCGACGACCACTGCTAAACAGAGTTAGCAGACGAGGATGTAGCGGCTGCTACAGAG 1109
Db 1797 ctggcagcagcactggttaacagattagcagagcgaggtatgtaggcggtgctcacagag 1856
QY 1110 TTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGGACAGTATTGGTATCTGCGCT 1169
Db 1857 ttcttgaagtgggtggcctaactacgctacactagaagagacagtattgtgtatctgcgct 1916
QY 1170 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACC 1229
Db 1917 ctgctgaagccagttaccttcggaagaaagagttgtagctcttgatccggaacaaacc 1976
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGGTGAAGCAGCAGAGTATACGCGCAGAAAAAGGA 1289
Db 1977 accgctggtagcggtgtttttttgttgaagcagcagagattacgctcagaaaaagga 2036
QY 1290 TCT 1292
Db 2037 tct 2039
RESULT 13
US-09-496-445-5
; Sequence 5, Application US/09496445
; GENERAL INFORMATION:
; APPLICANT: McNeish, John D.
; APPLICANT: Ahljinian, Michael K
; TITLE OF INVENTION: Transgenic Animals Expressing Human p25
; FILE REFERENCE: PC10142A
; CURRENT APPLICATION NUMBER: US/09/496,445
; CURRENT FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2462
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-496-445-5

Query Match 39.0%; Score 603; DB 18; Length 2462;
Best Local Similarity 100.0%; Pred. No. 1.1e-166;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTTGCTGGCGTTTTCATAGGCTCGCGCCCTCGACGAGCATCACAAAAATCGA 749
Db 359 gccgctgtgcgctgttttccataggtctcgcccccctgacgagcatcaaaaaatcga 418
QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACGAGCTATAAAGATACCGAGCGTTTCCCGCT 809
Db 419 cgctcaagtcagaggtgggaaacccgacagactataaagataccaggcggtttcccccct 478
QY 810 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCC 869
Db 479 ggaagctccctcgctgctctctgttccgacctgcttaaccgatacactgtccgcc 538
QY 870 TTCTTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTAGTAGTATCTCAGTTCG 929
Db 539 ttcttcccttcgggaagcggtggcgctttctcaatgctcacgctgtagtatactcagttcg 598
QY 930 GTGAGGTGCTGCTCCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGC 989
Db 599 gtgaggtgctgctcccaagctggcgctgtgtgcacgaaccccgcttcagcccgacccg 658
QY 990 TGGCCCTTATCCGCTAACTATCTGCTTGTAGTCCAAACCGGGTAAGACAGACTTATCGCCA 1049
Db 659 tggcccttataccgctaaactatctgttagtccaaacccggtaagacacgacttatacgcca 718
QY 1050 CTGCGACGACCACTGCTAAACAGAGTTAGCAGACGAGGATGTAGCGGCTGCTACAGAG 1109
Db 719 ctggcagcagcactggttaacagattagcagagcgaggtatgtaggcggtgctcacagag 778
QY 1110 TTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGGACAGTATTGGTATCTGCGCT 1169

Db 779 tttctgaagtgtggtcctaacctacggtctacactagaagagcagatttggatctgcgt 838
 QY 1170 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTGTAGCTCTTTCATCCGCAACAAACC 1229
 Db 839 ctgctgaagccagttaccttcggaagagctgtgtgctctgtatccgcgaacaaacc 898
 QY 1230 ACCGCTGTAGCGTGTGTTTTTGTTCGCAAGCAGCAGATTACGCGCAAGAAAGGA 1289
 Db 899 accgtgtgtagcgtgtgtttttttgttttgaagcagcagattacgcgcgaacaaaggga 958
 QY 1290 TCT 1292
 Db 959 tct 961
 RESULT 14
 US-60-118-478-5
 ; Sequence 5, Application US/60118478
 ; GENERAL INFORMATION:
 ; APPLICANT: McNeish, John D.
 ; APPLICANT: Ahlman, Michael K
 ; TITLE OF INVENTION: Transgenic Animals Expressing Human p25
 ; FILE REFERENCE: PC10142
 ; CURRENT APPLICATION NUMBER: US/60/118,478
 ; CURRENT FILING DATE: 1999-02-03
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0 - beta
 ; SEQ ID NO 5
 ; LENGTH: 2462
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-60-118-478-5

Query Match 39.0%; Score 603; DB 44; Length 2462;
 Best Local Similarity 100.0%; Pred. No. 1.1e-166;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 690 GCCGCGTGTGCGGTGTTTCCATAGCTCCGCCCTGACGAGCATCACAAAATCGA 749
 Db 359 gccgctgtgtggtgttttccatagctccgccctgacgagcatcacaaaatcga 418
 QY 750 CGCTCAAGTCAGAGTGGCGAAGCCGACGAGGACTATAAGATACGAGGTTTCCCTT 809
 Db 419 cgtcaagtcagaggtggcgaaacccgacagagactataagataccagcggttccct 478
 QY 810 GGAAGCTCCCTGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCC 869
 Db 479 ggaagctccctgtgcgtctctctgttccgacctgctccggtacacctgtccgc 538
 QY 870 TTTCTCCCTTCGGAAGCGTGGCGTTTCTCAATGCTACGCTGTAGGTATCTCAGTTCG 929
 Db 539 tttctcccttcggaagcgtggcgctttctcaatgtcagctgtaggtatctcagttcg 598
 QY 930 GTGTAGTGTGCTGCTCCAGTGGGCTGTGTCAGCAACCCCGTTCAGCCGACCGC 989
 Db 599 ggtgagtgctgtcgtcccaagctgggtgtgtgcaagaccccggttcagccgacgc 658
 QY 990 TCGCCTTATCCGGTAACTATCTGTTAGTCCAAACCGGTAAAGACAGCTATTCGCCA 1049
 Db 659 tgcgcttatccggttaactatctgttgcgtcccaacccggttaagacagactatcgca 718
 QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGAGGCGGTCTACAGAG 1109
 Db 719 ctggcagcagccactggttaacaggtattagcagagcaggtatgagcggtgctacagag 778
 QY 1110 TTTCTCAAGTGTGCGCTTAACCTACGCTTACCTAGAGCAGAGTATTTGGTATCTGCGCT 1169
 Db 779 tttctgaagtgtggtcctaacctacggtctacactagaagagcagatttggatctgcgt 838
 QY 1170 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTGTAGCTCTTTCATCCGCAACAAACC 1229

Db 839 ctgctgaagccagttaccttcggaagagctgtgtgctctgtatccgcgaacaaacc 898
 QY 1230 ACCGCTGTAGCGTGTGTTTTTGTTCGCAAGCAGCAGATTACGCGCAAGAAAGGA 1289
 Db 899 accgtgtgtagcgtgtgtttttgttttgaagcagcagattacgcgcgaacaaaggga 958
 QY 1290 TCT 1292
 Db 959 tct 961
 RESULT 15
 US-07-919-535C-26
 ; Sequence 26, Application US/07919535C
 ; GENERAL INFORMATION:
 ; APPLICANT: Haas, Werner
 ; APPLICANT: Hunziker, Willi
 ; TITLE OF INVENTION: SOLUBLE KIT LIGANDS
 ; NUMBER OF SEQUENCES: 68
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann-La Roche Inc.
 ; STREET: 340 Kingsland Street
 ; CITY: Nutley
 ; STATE: New Jersey
 ; COUNTRY: United States
 ; ZIP: 07110-1199
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: DOS 4.0
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/919,535C
 ; FILING DATE: 1993/07/23
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 91810609.7
 ; FILING DATE: 30-JUL-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kass, Alan P.
 ; REGISTRATION NUMBER: 32142
 ; REFERENCE/DOCKET NUMBER: RAN 4105/142
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (201) 235-4205
 ; TELEFAX: (201) 235-3500
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2577 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 121..126
 ; OTHER INFORMATION: /note= "Recognition site for
 ; OTHER INFORMATION: restriction enzyme NcoI"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 141..146
 ; OTHER INFORMATION: /note= "Recognition site for
 ; OTHER INFORMATION: restriction enzyme HindIII"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..2577
 ; OTHER INFORMATION: /note= "Complete nucleotide
 ; OTHER INFORMATION: sequence of E. coli
 ; OTHER INFORMATION: expression plasmid
 ; OTHER INFORMATION: POC 56/RBS II, NcoI"
 US-07-919-535C-26

Query Match		39.0%;	Score 603;	DB 3;	Length 2577;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
Best Local Similarity		100.0%;	Pred. No. 1.1e-166;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Matches 603;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
Qy	690	GC	CG	CG	CT			TG	CT	TG	CG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	CT	TG	

Search completed: January 17, 2002, 15:14:01
Job time: 22753 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:22 ; Search time 6501.33 Seconds
(without alignments)
2556.972 Million cell updates/sec

Title: US-09-242-202a-27

Perfect score: 1547

Sequence: 1 GGTACCTGCCACCATGGCGC.....CTGCACAGCGCTCTCCACACA 1547

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estfun:**

2: em_esthum:**

3: em_estin:**

4: em_estom:**

5: em_estpl:**

6: em_estba:**

7: em_estro:**

8: em_estov:**

9: em_htc:**

10: gb_estl:**

11: gb_est2:**

12: gb_htc:**

13: gb_gss:**

14: em_gss_fun:**

15: em_gss_hum:**

16: em_gss_inv:**

17: em_gss_pln:**

18: em_gss_pro:**

19: em_gss_tod:**

20: em_gss_vrt:**

21: em_gss_other:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597.2	38.6	1004	10	AJ281480
2	596.6	38.6	800	10	AJ281449
3	592.2	38.3	1070	10	AJ281552
4	585.4	37.8	756	10	BE749097
5	583.8	37.7	959	10	BE749147
6	580.4	37.5	759	10	BE749118
7	579.9	37.4	840	10	BE749178
8	576.2	37.2	784	13	AQ876119
9	573.6	37.1	795	13	AQ876011
10	565.2	36.5	617	10	AJ281661
11	555.2	35.9	571	10	AL044178
12	554	35.8	579	10	AJ281320

C	13	554	35.8	794	13	AQ875988
14	553.2	35.8	615	10	AQ875988	
15	550.2	35.6	954	10	AV735664	
16	549.6	35.5	1067	10	AL044364	
17	549.4	35.5	706	10	AU081137	
18	549.2	35.5	579	10	BE268532	
19	548.6	35.5	1045	11	AL043613	
20	548.2	35.4	629	10	AL043613	
21	547.2	35.4	568	10	BF664532	
22	546.2	35.3	700	10	AL593919	
23	544.8	35.2	982	10	AJ281376	
24	544.6	35.2	795	13	AJ281616	
25	543.2	35.1	548	10	BE749172	
26	543.2	35.1	718	13	AQ875935	
27	536.2	34.7	616	10	AJ281654	
28	532.2	34.4	615	10	AG010489	
29	529.2	34.2	857	13	AV735756	
30	528.6	34.2	747	10	AL044413	
31	524.4	33.9	890	11	AQ875006	
32	523	33.8	884	13	BE749123	
33	521.2	33.7	526	10	BG685100	
34	521.2	33.7	852	13	AQ491919	
35	518.2	33.5	1089	10	AL043840	
36	512.4	33.1	691	10	AL043840	
37	509.2	32.9	798	13	DKFZp434A	
38	507	32.8	636	10	DKFZp434A	
39	506.2	32.7	628	11	DKFZp434A	
40	505.4	32.7	902	10	DKFZp434A	
41	502.4	32.5	630	10	DKFZp434A	
42	500.4	32.3	608	13	DKFZp434G	
43	500.2	32.3	527	10	DKFZp434G	
44	499	32.3	799	13	DKFZp434G	
45	497.6	32.2	950	10	DKFZp434G	

ALIGNMENTS

RESULT	1
AJ281480	AJ281480
LOCUS	4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
DEFINITION	gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION	AJ281480
VERSION	AJ281480.1 GI:6929360
KEYWORDS	EST.
SOURCE	African malaria mosquito.
ORGANISM	Anopheles gambiae
REFERENCE	1 (bases 1 to 1004)
AUTHORS	Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE	20300950
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1. .1004 /organism="Anopheles gambiae" /strain="4A r/r" /db_xref="taxon:7165" /clone="4A3A-P4G8" /clone_lib="Anopheles gambiae immune competent 4A3A" /cell_line="immune competent 4A3A" /lab_host="E. coli DH10B"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN

Query Match 38.6%; Score 597.2; DB 10; Length 1004;
Best Local Similarity 99.3%; Pred. No. 1e-160;
Matches 599; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 690 GCCGGTGTGGCGCTTTTCATAGGCTCGCCGCCCTGACGAGCATCACAAAAATCGA 749
Db 24 GCCGGTGTGGCGCTTTTCATAGGCTCGCCGCCCTGACGAGCATCACAAAAATCGA 83

Qy 750 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTCGCCCT 809
Db 84 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTCGCCCT 143

Qy 810 GGAAGCTCCCTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACGTGTCGCC 869
Db 144 GGAAGCTCCCTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACGTGTCGCC 203

Qy 870 TTTCTCCCTCGGGAAGGTGGCGCTTCTCAATGCTCACGCTGAGTATCTCAAGTTCG 929
Db 204 TTTCTCCCTCGGGAAGGTGGCGCTTCTCATAGCTCACGCTGAGTATCTCAAGTTCG 263

Qy 930 GTGAGGTGCTTCCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGGTTACGCCGACCGC 989
Db 264 GTGAGGTGCTTCCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGGTTACGCCGACCGC 323

Qy 990 TGGCGCTTATCCGCTAATCTGCTTTCAGTCCACCCGCTTAAGACAGACTTATCGCCA 1049
Db 324 TGGCGCTTATCCGCTAATCTGCTTTCAGTCCACCCGCTTAAGACAGACTTATCGCCA 383

Qy 1050 CTGGCAGCAGCACTGGTAACAGGATTACGAGAGGAGTATGTAGCGGTGCTACAGAG 1109
Db 384 CTGGCAGCAGCACTGGTAACAGGATTACGAGAGGAGTATGTAGCGGTGCTACAGAG 443

Qy 1110 TCTTGAAGTGGTGGCTTAACAGGATTACGAGAGGAGTATGTAGTATCTGCGCT 1169
Db 444 TCTTGAAGTGGTGGCTTAACAGGATTACGAGAGGAGTATGTAGTATCTGCGCT 503

Qy 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTGTGTAGCTCTTGTATCCGCAACAAACC 1229
Db 504 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTGTGTAGCTCTTGTATCCGCAACAAACC 563

Qy 1230 ACCGCTGTAGCGGTGGTGTGTTTGTGTAAGCAGCAGATTAACCGCAGAAAAAGGA 1289
Db 564 ACCGCTGTAGCGGTGGTGTGTTTGTGTAAGCAGCAGATTAACCGCAGAAAAAGGA 623

Qy 1290 TCT 1292
Db 624 TCT 626

RESULT 2
AJ281449 800 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION AJ281449
VERSION AJ281449.1 GI:5929329
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea

; Anopheles.
1 (bases 1 to 800)
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. 800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4D5"
/cell_line="Anopheles immune competent 4A3A"
/lab_host="E. coli DH10b"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
203 a 206 c 198 g 193 t

Query Match 38.6%; Score 596.6; DB 10; Length 800;
Best Local Similarity 99.3%; Pred. No. 1.4e-160;
Matches 599; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 690 GCCGGTGTGGCGCTTTTCATAGGCTCGCCGCCCTGACGAGCATCACAAAAATCGA 749
Db 51 GCCGGTGTGGCGCTTTTCATAGGCTCGCCGCCCTGACGAGCATCACAAAAATAGA 110

Qy 750 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTCGCCCT 809
Db 111 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGCGGTTCGCCCT 170

Qy 810 GGAAGCTCCCTGCGCTCTCTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCC 869
Db 171 GGAAGCTCCCTGCGCTCTCTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCC 230

Qy 870 TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAAGCTCACGCTGAGTATCTCAAGTTCG 929
Db 231 TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGAGTATCTCAAGTTCG 290

Qy 930 GTGAGGTGCTGCTTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGC 989
Db 291 GTGAGGTGCTGCTTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGC 350

Qy 990 TGGCGCTTATCCGCTAATCTATCTGCTTTCAGTCCAAACCCGTAAGACAGACTTATCGCCA 1049
Db 351 TGGCGCTTATCCGCTAATCTATCTGCTTTCAGTCCAAACCCGTAAGACAGACTTATCGCCA 410

Qy 1050 CTGGCAGCAGCACTGGTAACAGGATTACGAGAGGAGTATGTAGCGGTGCTACAGAG 1109
Db 411 CTGGCAGCAGCACTGGTAACAGGATTACGAGAGGAGTATGTAGCGGTGCTACAGAG 470

Qy 1110 TCTTGAAGTGGTGGCTTAACAGGATTACGAGAGGAGTATGTAGTATCTGCGCT 1169
Db 471 TCTTGAAGTGGTGGCTTAACAGGATTACGAGAGGAGTATGTAGTATCTGCGCT 530

Qy 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTGTGTAGCTCTTGTATCCGCAACAAACC 1229
Db 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTGTGTAGCTCTTGTATCCGCAACAAACC 1229


```

Db 531 CTGCTGAAGCCAGTACTCTTCGGGAAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACCAACC 590
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTTTGAAGCAGCAGATTACGCCGACAGAAAAAGGA 1289
Db 591 ACCGCTGGTAGCGGTGTTTTTTTTTGAAGCAGCAGATTACGCCGACAGAAAAAGGA 650
QY 1290 TCT 1292
Db 651 TCT 653

RESULT 3
AJ281552
LOCUS
DEFINITION
4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION
AJ281552
VERSION
AJ281552.1 GI:6929432
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE
1 (bases 1 to 1070)
Dimopoulos G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
Location/Qualifiers
1..1070
/organism="Anopheles gambiae"
/strain="4A r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cdb="4A3A-P6F11"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT
263 a 283 c 255 g 269 t
ORIGIN

Query Match 38.3%; Score 592.2; DB 10; Length 1070;
Best Local Similarity 99.5%; Pred. No. 2.9e-159;
Matches 594; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 696 TTGCTGGGTTTTCATAGGCTCCGCCCTCAGCAGCATCACAAAATCGAGCTCA 755
Db 1 TTGCTGGGTTTTCATAGGCTCCGCCCTCAGCAGCATCACAAAATCGAGCTCA 60
QY 756 AGTCAGAGTGGCAAAACCCGACAGGACTATAAAGATACCAAGGGTTTCCCCCTGGAAGC 815
Db 61 AGTCAGAGTGGCAAAACCCGACAGGACTATAAAGATACCAAGGGTTTCCCTGGAAGC 120
QY 816 TCCCTGTGCGCTCTCTGTTCGACCGCTGCGGTACCGGATACCTGTGCGCTTCTC 875
Db 121 TCCCTGTGCGCTCTCTGTTCGACCGCTGCGGTACCGGATACCTGTGCGCTTCTC 180

```

```

QY 876 CTTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGGTGTAG 935
Db 181 CTTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGGTGTAG 240
QY 936 GTCGTTTCGCTCCAAAGCTGGGCTGTGTGACGAACCCCGTTCAGCGCGACCGTGGGCC 995
Db 241 GTCGTTTCGCTCCAAAGCTGGGCTGTGTGACGAACCCCGTTCAGCGCGACCGTGGGCC 300
QY 996 TTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGGTAAAGACACGACTTATCCGCACTGGCA 1055
Db 301 TTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGGTAAAGACACGACTTATCCGCACTGGCA 360
QY 1056 GCAGCCACTGTTAAGCAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTG 1115
Db 361 GCAGCCACTGTTAAGCAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTG 420
QY 1116 AAGTGGTGGCCTAACTACGGCTACACTAGAGGACAGATTGTTGGTATCTGCGCTCTGCTG 1175
Db 421 AAGTGGTGGCCTAACTACGGCTACACTAGAGGACAGATTGTTGGTATCTGCGCTCTGCTG 480
QY 1176 AAGCAGCTTACCTTCGGAAGAGTGGTAGCTCTTCGCGCAGAAAAAGAGTCT 1235
Db 481 AAGCAGCTTACCTTCGGAAGAGTGGTAGCTCTTCGCGCAGAAAAAGAGTCT 540
QY 1236 GGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAGAGTCT 1292
Db 541 GGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAAAGAGTCT 597

RESULT 4
BE749097/c
LOCUS
DEFINITION
601123138F1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348224 5',
mRNA sequence.
ACCESSION
BE749097
VERSION
BE749097.1 GI:10163089
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 756)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM140 row: k column: 09
High quality sequence start: 7
High quality sequence stop: 739.
FEATURES
Location/Qualifiers
1..756
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348224"
/cdb="NIH_MGC_5"
/tissue_type="carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: cervix; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

BASE COUNT 184 a 191 c 207 g 174 t

ORIGIN

Query Match 37.8%; Score 585.4; DB 10; Length 756;
Best Local Similarity 97.3%; Pred. No. 2.4e-157;
Matches 606; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 671 GAAAGAAATTCCTGCACGCCGCGCTGTGGCG-TTTTTCATAGCTCCGCCCCCTG 729
Db 738 GACCGCTTTTGTGTGATCCCGTGTGGCGTTTTTTCATAGCTCCGCCCCCTG 679

QY 730 ACAGCATCACAAAATACAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAA 789
Db 678 ACAGCATCACAAAATACAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAA 619

QY 790 GATACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCGACCTGCCG 849
Db 618 GATACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCGAGCTGCCG 559

QY 850 TTACCGGATACCTGTCGCGCTTTCTCCCTCGGAAAGCGTGGCGTTTCTCAATGCTAC 909
Db 558 TTACCGGATACCTGTCGCGCTTTCTCCCTCGGAAAGCGTGGCGTTTCTCATAGCTAC 499

QY 910 GCTGTAGTATCTCAGTTCGGTGTAGTTCGTCGCTCCAAAGCTGGCTGTGTGCACGAAC 969
Db 498 GCTGTAGTATCTCAGTTCGGTGTAGTTCGTCGCTCCAAAGCTGGCTGTGTGCACGAAC 439

QY 970 CCCCCCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTTCAGTCCAAACCCG 1029
Db 438 CCCCCCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTTCAGTCCAAACCCG 379

QY 1030 TAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT 1089
Db 378 TAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT 319

QY 1090 ATGTAGCGGTGTCACAGAGTTCCTGAAAGTGGGCTTAACCTACGCTACACTAGAAGGA 1149
Db 318 ATGTAGCGGTGTCACAGAGTTCCTGAAAGTGGGCTTAACCTACGCTACACTAGAAGGA 259

QY 1150 CAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCT 1209
Db 258 CAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCT 199

QY 1210 CTTGATCCGGCAAAACAAACCCAGCGTGGTAGCGGTGGTTTTTTTGTGCAACGACAGA 1269
Db 198 CTTGATCCGGCAAAACAAACCCAGCGTGGTAGCGGTGGTTTTTTTGTGCAACGACAGA 139

QY 1270 TTACGGCGAGAAAAAGGATCT 1292
Db 138 TTACGGCGAGAAAAAGGATCT 116

RESULT 5
BE749147/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE749147 959 bp mRNA EST 15-SEP-2000
60112315F1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348291 5',
mRNA sequence.
BE749147
BE749147.1 GI:10163139
EST.
human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 959)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC#140 row: n column: 04
High quality sequence start: 13
High quality sequence stop: 800.
Location/Qualifiers
1. .959
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348291"
/clone_lib="NIH_MGC_5"
/tissue_type="carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: cervix; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life technologies)."

BASE COUNT 245 a 229 c 252 g 233 t
ORIGIN

Query Match 37.7%; Score 583.8; DB 10; Length 959;
Best Local Similarity 97.1%; Pred. No. 7.4e-157;
Matches 605; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 671 GAAAGAAATTCCTGCACGCCGCGCTGTGGCGTT-TTTCATAGCTCCGCCCCCTG 729
Db 731 GACCGCTTTTGTGTGATCCCGTGTGGCTGTTTTCATAGCTCCGCCCCCTG 672

QY 730 ACAGCATCACAAAATACAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAA 789
Db 671 ACAGCATCACAAAATACAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAA 612

QY 790 GATACAGCGTTTCCCTCGGAAAGCTCCCTCGTGGCTCTCTGTTCGACCTGCCG 849
Db 611 GATACAGCGTTTCCCTCGGAAAGCTCCCTCGTGGCTCTCTGTTCGACCTGCCG 552

QY 850 TTACCGGATACCTGTCGCGCTTTCTCCCTCGGAAAGCGTGGCGTTTCTCAATGCTAC 909
Db 551 TTACCGGATACCTGTCGCGCTTTCTCCCTCGGAAAGCGTGGCGTTTCTCATAGCTAC 492

QY 910 GCTGTAGTATCTCAGTTCGGTGTAGTTCGTCGCTCCAAAGCTGGGCTGTGTGCACGAAC 969
Db 491 GCTGTAGTATCTCAGTTCGGTGTAGTTCGTCGCTCCAAAGCTGGGCTGTGTGCACGAAC 432

QY 970 CCCCCCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTTCAGTCCAAACCCG 1029
Db 431 CCCCCCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTTCAGTCCAAACCCG 372

QY 1030 TAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT 1089
Db 371 TAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT 312

QY 1090 ATGTAGCGGTGTCACAGAGTTCCTGAAAGTGGGCTTAACCTACGCTACACTAGAAGGA 1149
Db 311 ATGTAGCGGTGTCACAGAGTTCCTGAAAGTGGGCTTAACCTACGCTACACTAGAAGGA 252

QY 1150 CAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCT 1209
Db 251 CAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCT 192

QY 1210 CTTGATCCGGCAAAACAAACCCAGCGTGGTAGCGGTGGTTTTTTTGTGCAACGACAGA 1269
Db 191 CTTGATCCGGCAAAACAAACCCAGCGTGGTAGCGGTGGTTTTTTTGTGCAACGACAGA 132

QY 1270 TTACGGCGAGAAAAAGGATCT 1292
|||||

```

Db 131 TTACGGCCAGAAAAAGGATCT 109

RESULT 6
BE749118/c
LOCUS BE749118.1 759 bp mRNA EST 15-SEP-2000
DEFINITION 601123194F1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348238 5',
mRNA sequence.
ACCESSION BE749118
VERSION BE749118.1 GI:10163110
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 759)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM140 row: k column: 23
High quality sequence start: 8
High quality sequence stop: 747.
Location/Qualifiers
1. 759
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348238"
/tissue_type="carcinoma cell line"
/lab_host="NIH_MGC_5"
/note="Organ: cervix; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 191 a 193 c 203 g 172 t
ORIGIN
Query Match 37.5%; Score 580.4; DB 10; Length 759;
Best Local Similarity 98.7%; Pred. No. 6.6e-156;
Matches 596; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 691 CCGCGTTGCTGGCTTTTCATAGGCTCCGCCCTCGACGAGCATC--ACAAAAATCG 748
Db 734 CCGCGTTGCTGGCTTTTCATAGGCTCCGCCCTCGACGAGCATCTACTAATATCG 675
QY 749 ACGCTCAAGTCAGAGTGGCGAAACCGACAGAGACTATAAGATACACGAGCGTTTCCGCC 808
Db 674 ACGCTCAAGTCAGAGTGGCGAAACCGACAGAGACTATAAGATACACGAGCGTTTCCGCC 615
QY 809 TGGAGGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGC 868
Db 614 TGGAGGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGC 555
QY 869 CTTTCTCCCTTCGGAGCGTGGCGCTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTC 928
Db 554 CTTTCTCCCTTCGGAGCGTGGCGCTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTC 495
QY 929 GGTGTAGGTCGTTGCTGCCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCGACCG 988
Db 494 GGTGTAGGTCGTTGCTGCCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCGACCG 435

```

```

QY 989 CTGCGCCTTATCCGGTAACATATCGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCC 1048
Db 434 CTGCGCCTTATCCGGTAACATATCGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCC 375
QY 1049 ACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTTACAGA 1108
Db 374 ACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTTACAGA 315
QY 1109 GTTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGC 1168
Db 314 GTTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGC 255
QY 1169 TCTGCTGAAGCAGTACCTTCGGAAGAAAGAGTTGGTAGCTCTTGTATCGCGCAAAACAAC 1228
Db 254 TCTGCTGAAGCAGTACCTTCGGAAGAAAGAGTTGGTAGCTCTTGTATCGCGCAAAACAAC 195
QY 1229 CACCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGAGATTACGCGCAGAAAAAAGG 1288
Db 194 CACCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGAGATTACGCGCAGAAAAAAGG 135
QY 1289 ATCT 1292
Db 134 ATCT 131

RESULT 7
BE749178/c
LOCUS BE749178.1 840 bp mRNA EST 15-SEP-2000
DEFINITION 601123444F1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348154 5',
mRNA sequence.
ACCESSION BE749178
VERSION BE749178.1 GI:10163170
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM140 row: h column: 11
High quality sequence start: 4
High quality sequence stop: 782.
Location/Qualifiers
1. 840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348154"
/tissue_type="carcinoma cell line"
/lab_host="NIH_MGC_5"
/note="Organ: cervix; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 210 a 205 c 225 g 200 t
ORIGIN

```

Query Match		37.4%; Score 579; DB 10; Length 840;
Best Local Similarity		99.1%; Pred. No. 1.7e-155;
Matches		582; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	706	TTTTTCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACCTCAAGTCAGAGT 765
Db	750	TTTTTCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACCTCAAGTCAGAGT 691
QY	766	GGCGAAACCGACGAGCTATAAGATACACAGCGCTTTCCTCCCTGGAAGTCCCTCGTGC 825
Db	690	GGCGAAACCGACGAGCTATAAGATACACAGCGCTTTCCTCCCTGGAAGTCCCTCGTGC 631
QY	826	GCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGGA 885
Db	630	GCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGGA 571
QY	886	GCCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTTGCCT 945
Db	570	GCCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTTGCCT 511
QY	946	CCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTA 1005
Db	510	CCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTA 451
QY	1006	ACTATCGTCTTGAGTCCAAACCCCGTAAAGACACACTTATGCGCACTGGCAGCCACTG 1065
Db	450	ACTATCGTCTTGAGTCCAAACCCCGTAAAGACACACTTATGCGCACTGGCAGCCACTG 391
QY	1066	GTAACAGGATTACAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGC 1125
Db	390	GTAACAGGATTACAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGC 331
QY	1126	CTAACTACGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTA 1185
Db	330	CTAACTACGGCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTA 271
QY	1186	CTTCGGAAGAGTGGTGTAGTCTTTGATCGCGCAAAACACCCGCTGTGAGCGGTG 1245
Db	270	CTTCGGAAGAGTGGTGTAGTCTTTGATCGCGCAAAACACCCGCTGTGAGCGGTG 211
QY	1246	GTTTTTTTGTTTTCAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 1292
Db	210	GTTTTTTTGTTTTCAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 164
RESULT 8		
A0876119/c		
LOCUS		
DEFINITION		V133El mtm-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
ACCESSION		A0876119
VERSION		A0876119.1 GI:6288363
KEYWORDS		GSS.
SOURCE		baker's yeast.
ORGANISM		Saccharomyces cerevisiae
REFERENCE		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS		Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
		1 (bases 1 to 784)
		Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,
		desEtages, S. A., Cheung, K. H., Sheehan, A., Symoniat, D., Jansen, R.,
		Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
		Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
		Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
		Gene Disruption
		Unpublished (1999)
TITLE		Contact: Kumar A
JOURNAL		Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
COMMENT		Yale University
		P.O. Box 208103, New Haven, CT 06520-8103, USA
		Tel: 203 432 9949
		Fax: 203 432 6161
		Email: anuj.kumaryale.edu
		te of mtm-3xHA/lacZ insertion.
Seq primer: GGCCTCTTTCTTTTGGAGTAC		
Class: transposon-tagged.		
Location/Qualifiers		
1. .784		
/organism="Saccharomyces cerevisiae"		
/strain="Y2278 - S288C background, cir(0) rho(0)"		
/db_xref="taxon:4932"		
/clone_lib="mtm-3xHA/lacZ Insertion Library, strain Y2278"		
/lab_host="E. coli"		
/note="Vector: pHS6-Sal; A yeast genomic DNA library		
without 2 micron or mitochondrial DNA was prepared in		
pHS6-Sal; genomic DNA was size-fractionated (DNA of		
roughly 2-3 kb in length) prior to cloning. This library		
was subsequently mutagenized with a mtm-3xHA/lacZ		
minitransposon containing lacZ, URA3, and tet resistance.		
BASE COUNT		187 a 200 c 210 g 186 t 1 others
ORIGIN		
Query Match		37.2%; Score 576.2; DB 13; Length 784;
Best Local Similarity		97.7%; Pred. No. 1.1e-154;
Matches		584; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY	695	GTGTGCTGGCGTTCATAGGCTCCGCCCTGACGAGCATCACAAAATCGAGCTC 754
Db	784	GTGTGCTGGCGTTCATAGGCTCCGCCCTGACGAGCATCACAAAATCGAGCTC 725
QY	755	AAGTCAGAGTGGCGAAACCCGACAGACTATAAGATACACAGCGCTTTCCTCCCTGGAAG 814
Db	724	AAGTCAGAGTGGCGAAACCCGACAGACTATAAGATACACAGCGCTTTCCTCCCTGGAAG 665
QY	815	CTCCCTCGTGGCTCTCCTGTTCCGACCTTACCGGATACCTGTCGCGCTTTCT 874
Db	664	CTCCCTCGTGGCTCTCCTGTTCCGACCTTACCGGATACCTGTCGCGCTTTCT 605
QY	875	CCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGA 934
Db	604	CCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGA 545
QY	935	GGTCGTTCGCTCCAAAGCTGGGCTGTGTCAGCAACCCCGCTTCAGCCGACCTGCGC 994
Db	544	GGTCGTTCGCTCCAAAGCTGGGCTGTGTCAGCAACCCCGCTTCAGCCGACCTGCGC 485
QY	995	CTTATCCGCTTAATCTGCTTTCAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGGC 1054
Db	484	CTTATCCGCTTAATCTGCTTTCAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGGC 425
QY	1055	AGCAGCCTGTTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT 1114
Db	424	AGCAGCCTGTTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT 365
QY	1115	GAAGTGTGGCTTAACCTACGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCT 1174
Db	364	GAAGTGTGGCTTAACCTACGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCT 305
QY	1175	GAAGCCAGTTTACCTTCGGAAGAAAGTGGTGTAGTCTTGTATCCGCGCAACAAACACCGC 1234
Db	304	GAAGCCAGTTTACCTTCGGAAGAAAGTGGTGTAGTCTTGTATCCGCGCAACAAACACCGC 245
QY	1235	TGTTAGCGGTGGTTTTTTTTCGTCAGAGCAGAGATTACCGCGCAGAAAAAAGGATCT 1292
Db	244	TGTTAGCGGTGGTTTTTTTTCGTCAGAGCAGAGATTACCGCGCAGAAAAAAGGATCT 187
RESULT 9		
A0876011/c		
LOCUS		
DEFINITION		V132B5 mtm-3xHA/lacZ Insertion Library, strain Y2278 Saccharomycetes
ACCESSION		A0876011
VERSION		A0876011.1 GI:6288255
KEYWORDS		GSS.

SOURCE	baker's yeast.
ORGANISM	Saccharomyces cerevisiae
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
AUTHORS	1 (bases 1 to 795) Ross-Macdonald,P.,Roemer,T.,Coelho,P.S.R.,Agarwal,S.,Kumar,A.,desEtages,S.A.,Cheung,K.H.,Sheehan,A.,Symonistis,D.,Jansen,R.,Umansky,L.,Heidtmann,M.,Nelson,K.,Iwasaki,H.,Kanada,D.,Lugo,R.,Hager,K.,Miller,P.,Roeder,G.S. and Snyder,M.
TITLE	Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
JOURNAL	Unpublished (1999)
COMMENT	Contact: Kumar A Michael Snyder, Dept. of Mol. Cell. and Dev. Biology Yale University P.O. Box 208103, New Haven, CT 06520-8103, USA Tel: 203 432 9949 Fax: 203 432 6161 Email: anuj.kumar@yale.edu te of mtm-3xHA/lacZ insertion. Seq primer: GGCGTCTCTTTCTTTGGAGTAC Class: transposon-tagged.
FEATURES	Location/Qualifiers
source	1..795 /organism="Saccharomyces cerevisiae" /strain="Y27878 - S288C background, clr(0) rho(0)" /db_xref="taxon:4932" /clone_lib="mtm-3xHA/lacZ Insertion Library, strain Y27878" /lab_host="E. coli" /note="Vector: pHS56-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtm-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT	191 a 203 c 212 g 188 t 1 others
ORIGIN	
Query Match	37.1%; Score 573.6; DB 13; Length 795;
Best Local Similarity	99.1%; Pred. No. 6e-154;
Matches 576; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
QY	712 ATAGGCTCCGCCCCCTGACGAGCATCACAAAAATGACGCTCAAGTCAGAGGTGGCGAA 771
Db	795 ATAGGCTCCGCCCCCTGACGAGCATCACAAAAATGACGCTCAAGTCAGAGGTGGCGAA 736
QY	772 ACCGACAGGACTATAAGATACACAGGGTTCCGCCCTGGAAGTCCCTCGTGGCTCTC 831
Db	735 ACCGACAGGACTATAAGATCCACAGGGTTCCGCCCTGGAAGTCCCTCGTGGCTCTC 676
QY	832 CTGTTCCGACCTCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGACGCGGG 891
Db	675 CTGTTCCGACCTCCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGACGCGGG 616
QY	892 CGCTTTCTCAATGCTACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCCAAGC 951
Db	615 CGCTTTCTCAGTCTACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCCAAGC 556
QY	952 TGGGCTGTGTCAGACACCCCGCTTCAGCCCGACCGCTGGCGCTTATCCGGTAACTATC 1011
Db	555 TGGGCTGTGTCAGAAACCCNCCGTTACGCCCGACCGCTGGCGCTTATCCGGTAACTATC 496
QY	1012 GTCTTGAGTCCAAACCCGTRAGACACGACTTATCGCCACTGGCAGCACCCTGTTAACA 1071
Db	495 GTCTTGAGTCCAAACCCGTRAGACACGACTTATCGCCACTGGCAGCACCCTGTTAACA 436
QY	1072 GGATTACGACGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTGAAGTGGTGGCCCTAACT 1131
Db	435 GGATTACGACGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTGAAGTGGTGGCCCTAACT 376
QY	1132 ACGSCTACACTAGAAGGACACAGTATTTTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTCG 1191

Db	375	ACGGCTACACTAGAAGACAGTATTTGGTATCTGCTGCTGAAGCCAGTACCTTCG	316
Qy	1192	GAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCGCTGGTAGCGGTGTTTT	1251
Db	315	GAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCAGCGCTGGTAGCGGTGTTTT	256
Qy	1252	TTGTTTGAAGCAGCAGGATACGGCGCAGAAAAAAGGATCT	1292
Db	255	TTGTTTGAAGCAGCAGGATACGGCGCAGAAAAAAGGATCT	215
RESULT 10			
AJ281661			
LOCUS			
DEFINITION			
ACCESSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	690	GCCGCTGCTGGCGTTTTTCATAGGCTCGCGCCCTGACGAGCATCAAAAAATCGA	749
Db	37	GCCGCTGCTGGCGTTTTTCATAGGCTCGCGCCCTGACGAGCATCAAAAAATCGA	96
Qy	750	CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCGAGCGTTCCCGCCT	809
Db	97	CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCGAGCGTTCCCGCCT	156
Qy	810	GGAAAGTCCTCTGCGCTCTCTCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCGCC	869

```
Db 157 GGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTTACCGGATACCTGTCCGGC 216
QY 870 TTTCTCCCTTCGGGAAGCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCC 929
Db 217 TTTCTCCCTTCGGGAAGCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCC 276
QY 930 GTCTAGTGTCTCGCTCCCAAGCTGGCTGTGTGACGAACCCCGCTTTCAGCCGACCCG 989
Db 277 GTGTAGTGTCTCGCTCCCAAGCTGGCTGTGTGACGAACCCCGCTTTCAGCCGACCCG 336
QY 990 TGGCCCTTATCCGGTAACATATCGTCTTGAGTCCCAACCCCGTAAGACACGACTTATCGCCA 1049
Db 337 TGGCCCTTATCCGGTAACATATCGTCTTGAGTCCCAACCCCGTAAGACACGACTTATCGCCA 396
QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTACGAGCAGGAGTATGTAGGCGGTGCTACAGAG 1109
Db 397 CTGGCAGCAGCCACTGGTAACAGGATTACGAGCAGGAGTATGTAGGCGGTGCTACAGAG 456
QY 1110 TTTCTCAAGTGGTGGCTTAACAGGCTTACAGGCTACACTAGAGGACAGTATTTGGTATCTCGCT 1169
Db 457 TTTCTCAAGTGGTGGCTTAACAGGCTTACAGGCTACACTAGAGGACAGTATTTGGTATCTCGCT 516
QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTTTGTATCCGCGCAACAAACC 1229
Db 517 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTTTGTATCCGCGCAACAAACC 576
QY 1230 ACCGCTGGTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCAGATT 1271
Db 577 ACCGCTGGTAGC-GTGGTGTGTTGTTGTTGCAAGCAGCAGATT 617

RESULT 11
AL044178 571 bp mRNA EST 29-FEB-2000
LOCUS DKF2p434P0828_s1.434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKF2p434P0828 3', mRNA sequence.
ACCESSION AL044178
VERSION AL044178.3 GI:5935954
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Bloecker, et al.)
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5866789.
Contact: Bloecker H
MIPS
Am Klopferpitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GfR (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKF2p434P0828) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434P0828"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH108"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

FEATURES
SOURCE
122 a 168 c 149 g 132 t

BASE COUNT
```

ORIGIN

Query Match 35.9%; Score 555.2; DB 10; Length 571;
Best Local Similarity 99.5%; Pred. No. 1.1e-148;
Matches 557; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 690 GCCGGTGTGCTGGCGCTTTTCCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 749
Db 12 GCCGGTGTGCTGGCGCTTTTCCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 71
QY 750 CGCTCAAGTFCAGAGTGGGGAACCCGACAGGACTATAAGATACACGGCGTTTCCCGCT 809
Db 72 CGCTCAAGTFCAGAGTGGGGAACCCGACAGGACTATAAGATACACGGCGTTTCCCGCT 131
QY 810 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTTACCGGATACCTGTCCGCG 869
Db 132 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTTACCGGATACCTGTCCGCG 191
QY 870 TTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCC 929
Db 192 TTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCC 251
QY 930 GTGTAGTGTCTCGCTCCAAAGCTGGCTGTGTGACGAACCCCGCTTTCAGCCGACCCG 989
Db 252 GTGTAGTGTCTCGCTCCAAAGCTGGCTGTGTGACGAACCCCGCTTTCAGCCGACCCG 311
QY 990 TCGCCCTTATCCGGTAACATATCGTCTTGAGTCCCAACCCCGTAAGACACGACTTATCGCCA 1049
Db 312 TCGCCCTTATCCGGTAACATATCGTCTTGAGTCCCAACCCCGTAAGACACGACTTATCGCCA 371
QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTACGAGCAGGAGTATGTAGGCGGTGCTACAGAG 1109
Db 372 CTGGCAGCAGCCACTGGTAACAGGATTACGAGCAGGAGTATGTAGGCGGTGCTACAGAG 431
QY 1110 TTTCTCAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGCT 1169
Db 432 TTTCTCAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGCT 491
QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTTTGTATCCGCGCAACAAACC 1229
Db 492 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTTTGTATCCGCGCAACAAACC 551
QY 1230 ACCGCTGGTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCAGATT 1249
Db 552 ACCGCTGGTAGCGGTGGTGTGTTTGTGTTGCAAGCAGCAGATT 571
```

RESULT 12

```
AJ281320 579 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-PIH1-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-PIH1, mRNA sequence.
ACCESSION AJ281320
VERSION AJ281320.1 GI:6929201
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles
1 (bases 1 to 579)
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
```

FEATURES	Meyerothofstrasse 1, 69117 Heidelberg, Germany.	
	Location/Qualifiers	
source	1. .579	
	/organism="Anopheles gambiae"	
REFERENCE	/strain="4A r/r"	
	/db_xref="taxon:7165"	
AUTHORS	/clone="4A3A-PH1"	
	/cell_line="immune competent 4A3A"	
TITLE	/lab_host="E. coli DH10B"	
	/note="Vector: pT773D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."	
JOURNAL	120 a 167 c 151 g 137 t	
	4 others	
COMMENT	BASE COUNT	
	120 a 167 c 151 g 137 t	
FEATURES	Query Match 35.8%; Score 554; DB 10; Length 579;	
	Best Local Similarity 98.1%; Pred. No. 2.4e-148;	
source	Matches 557; Conservative 3; Mismatches 8; Indels 0; Gaps 0;	
	QY 690 GCGCGTGTGCGGTTTTCATAGCTCCGCCCTGACGAGCATCAAAAATCGA 749	
source	Db 5 GCGCGTGTGCGGTTTTCATAGCTCCGCCCTGACGAGCATCAAAAATCGA 64	
	QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCGTTTCCCGCT 809	
source	Db 65 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCGTTTCCCGCT 124	
	QY 810 GGAAGTCCCTCGTGGCTCTCTGTTTCGACCGCTGCGGTTACCGGATACCTGTCCGCC 869	
source	Db 125 GGAAGTCCCTCGTGGCTCTCTGTTTCGACCGCTGCGGTTACCGGATACCTGTCCGCC 184	
	QY 870 TTTCTCCCTTCGGGAGCGTGGCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTCG 929	
source	Db 185 TTTCTCCCTTCGGGAGCGTGGCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTCG 244	
	QY 930 GTGTAGTCTGCTCCGCTCCAAAGCTGGGCTGTGTGCAGCAACCCCGCTTACGCCGACCGC 989	
source	Db 245 GTGTAGTCTGCTCCGCTCCAAAGCTGGGCTGTGTGCAGCAACCCCGCTTACGCCGACCGC 304	
	QY 990 TGGCGCTTATCCGCTAACTATCGCTTTAGTCCAAACCCGGTAAGACACGACTTATCGCCA 1049	
source	Db 305 TGGCGCTTATCCGCTAACTATCGCTTTAGTCCAAACCCGGTAAGACACGACTTATCGCCA 364	
	QY 1050 CTGGCAGCGCTGTATACAGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAG 1109	
source	Db 365 CTGGCAGCGCTGTATACAGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAG 424	
	QY 1110 TTTCTGAAGTGGTGGCTTAACCTACCGGCTACACTAGAGGACAGTATTTGTTATCTCGCT 1169	
source	Db 425 TTTCTGAAGTGGTGGCTTAACCTACCGGCTACACTAGAGGACAGTATTTGTTATCTCGCT 484	
	QY 1170 CTGCTGAAGCCAGTATACCTTCGGAAGAAAGTTGGTAGCTCTTTGATCCGCAACAAACC 1229	
source	Db 485 CTGCTGAAGCCAGTATACCTTCGGAAGAAAGTTGGTAGCTCTTTGATCCGCAACAAACC 544	
	QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGT 1257	
source	Db 545 ACCGCTGGTAGCGGTGTTTTTTTGT 572	
RESULT	13	
	AQ875988/c	
LOCUS	794 bp DNA	
	08-NOV-1999	
DEFINITION	V131h3 mtN-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', DNA sequence.	
	AQ875988	
ACCESSION	13	
	AQ875988/c	
VERSION	AQ875988.1 GI:6288232	
	GSS.	
KEYWORDS	baker's yeast.	
	Saccharomyces cerevisiae	
SOURCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.	
ORGANISM	1 (bases 1 to 794)	
	Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,	
REFERENCE	desRages, S. A., Cheung, K. H., Sheehan, A., Symoniatidis, D., Jansen, R.,	
	Mansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,	
AUTHORS	Hager, K., Miller, P., Roeder, G. S., and Snyder, M.	
	Large-Scale Analysis of the Yeast Genome by Transposon Tagging and	
TITLE	Gene Disruption	
	Unpublished (1999)	
JOURNAL	Contact: Kumar A	
	Michael Snyder, Dept. of Mol. Cell. and Dev. Biology	
COMMENT	Yale University	
	P.O. Box 208103, New Haven, CT 06520-8103, USA	
FEATURES	Tel: 203 432 9949	
	Fax: 203 432 6161	
source	Email: anuj.kumar@yale.edu	
	te of mtN-3xHA/lacZ insertion.	
FEATURES	Seq primer: GGCCTCTTCTTTGGAAGTAC	
	Class: transposon-tagged.	
source	Location/Qualifiers	
	1. .794	
FEATURES	/organism="Saccharomyces cerevisiae"	
	/strain="Y2278 - S288C background, cir(0) rho(0)"	
source	/db_xref="taxon:4932"	
	/clone_lib="mtN-3xHA/lacZ Insertion Library, strain Y2278"	
FEATURES	/lab_host="E. coli"	
	/note="Vector: pHS56-Sal; A yeast genomic DNA library	
source	without 2 micron or mitochondrial DNA was prepared in	
	pHS56-Sal; genomic DNA was size-fractionated (DNA of	
FEATURES	roughly 2-3 kb in length) prior to cloning. This library	
	was subsequently mutagenized with a mtN-3xHA/lacZ	
source	minitransposon containing lacZ, URA3, and tet resistance.	
	BASE COUNT	
ORIGIN	186 a 193 c 220 g 191 t 4 others	
	Query Match 35.8%; Score 554; DB 13; Length 794;	
source	Best Local Similarity 98.3%; Pred. No. 2.6e-148;	
	Matches 579; Conservative 0; Mismatches 8; Indels 2; Gaps 2;	
source	QY 706 TTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGAGTCAGTCAGAGGT 765	
	Db 788 TTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGAGTCAGTCAGAGGT 729	
source	QY 766 GCGCAACCGGACA-GGACTATAAAGATA-CCAGGCGTTTCCCGCTGGAAGCTCCCTCGT 823	
	Db 728 GCGCAACCGGACAAGGCGCTATAAAGATATCCAGGCGTTTCCCGCTGGAAGCTCCCTCGT 669	
source	QY 824 GCGCTCTCTGTTCCGACCTGCGCTTACGGATACCTGTCGCCCTTCTCCCTTCGGG 883	
	Db 668 GCGCTCTCTGTTCCGACCTGCGCTTCCGATATCTGTCGCCCTTCTCCCTTCGGG 609	
source	QY 884 AAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGTGTTCG 943	
	Db 608 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGTGTTCG 549	
source	QY 944 CTCGAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGGCTTATCCGG 1003	
	Db 548 CTCGAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGGCTTATCCGG 489	
source	QY 1004 TAAGTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCGGCACTGGCAGCAGCAC 1063	
	Db 488 TAAGTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCGGCACTGGCAGCAGCAC 429	
source	QY 1064 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGCGGTCTACAGAGTCTTTGAAGTGGTG 1123	
	Db 428 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGCGGTCTACAGAGTCTTTGAAGTGGTG 369	

```
QY 1124 GCCTAACTACGGCTACACTAGAGGACACGATTTTGGTATCTGCGTCTCTGCTGAAGCCAGT 1183
|||||
Db 368 GCCTAACTACGGCTACACTAGAGGACACGATTTTGGTATCTGCGTCTCTGCTGAAGCCAGT 309
|||||
QY 1184 TACCTTCGGAAGAGGTTGGTAGCTCTTTGATTCGCGCAACAAACCCGCTGGTAGCGG 1243
|||||
Db 308 TACCTTCGGAAGAGGTTGGTAGCTCTTTGATTCGCGCAACAAACCCGCTGGTAGCGG 249
|||||
QY 1244 TGGTTTTTTTGTTCGAACGACGATTTACGCGCGCAAAAAAAGGATCT 1292
|||||
Db 248 TGGTTTTTTTGTTCGAACGACGATTTACGCGCGCAAAAAAAGGATCT 200
|||||

RESULT 14
AV735664 AV735664 615 bp mRNA EST 17-OCT-2000
LOCUS
DEFINITION AV735664 CB Homo sapiens cDNA clone CBNAME07 5', mRNA sequence.
ACCESSION AV735664
VERSION AV735664.1 GI:10853245
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Zhang, Q., Ye, M., Wu, X., Gu, J., Huang, Q., Zhou, J., Shen, Y., Han, Z.,
Chen, S., Mao, M. and Chen, Z.
Homo sapiens CB library cDNA clones
Unpublished (2000)
CONTACT: Zhu Chen
Shanghai Institute of Hematology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64740490
Fax: 86-21-64743206
Email: mbshi@ms.scn.sh.cn
This clone is available at Shanghai Hematology Institute in
Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
Location/Qualifiers
FEATURES
source
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CBNAME07"
/clone_lib="CB"
/tissue_type="cord blood"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/notes="Vector: pBluescript; Site_1: EcoRI; The insert is
cloned randomly with the EcoRI digestion"
BASE COUNT 129 a 176 c 159 g 147 t
ORIGIN
Query Match 35.8%; Score 553.2; DB 10; Length 615;
Best Local Similarity 97.5%; Pred. No. 4.1e-148;
Matches 591; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 688 CGCGCGCGTTGCTGGGCTTTTCCATAGGCTCGCGCCCTGACGAGCATCAAAAATC 747
|||||
Db 7 CGCGCGCGTTGCTGGGCTTTTCCATAGGCTCGCGCCCTGACGAGCATCAAAAATC 66
|||||
QY 748 GAGCGTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCC 807
|||||
Db 67 GAGCGTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCC 126
|||||
QY 808 CTGGAAGCTCCCTCGTGGCGCTCTCTCTGTTCCGACCGCTGCCGTTACCGGATACCTGTCCG 867
|||||
Db 127 CTGGAAGCTCCCTCGTGGCGCTCTCTCTGTTCCGACCGCTGCCGTTACCGGATACCTGTCCG 186
|||||
QY 868 CCTTTCTCCCTTCGGGAACGCTGGCGCTTTCATGCTCAGCGCTAGTATCTCAGTT 927
|||||
```

```
Db 187 CCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTAGGTATCTCAGTT 246
QY 928 CGGTGTAGGTCTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACC 987
|||||
Db 247 CGGTGTAGGTCTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACC 306
|||||
QY 988 GCTGGCGCTTATCCGGTAACCTATCTGTTAGTTCACCAACCCGGTAAGACACGACTTATCCG 1047
|||||
Db 307 GCTGGCGCTTATCCGGTAACCTATCTGTTAGTTCACCAACCCGGTAAGACACGACTTATCCG 366
|||||
QY 1048 CACTGGCAGCAGCCACTGTGTACAGATTTAGCAGCGAGGTATGTAGCGGTGTCTACAG 1107
|||||
Db 367 CACTGGCAGCAGCCACTGTGTACAGATTTAGCAGCGAGGTATGTAGCGGTGTCTACAG 426
|||||
QY 1108 AGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGG 1167
|||||
Db 427 AGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGG 486
|||||
QY 1168 CTCTGCTGAAGCCAGTTACC-TTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAA 1226
|||||
Db 487 CTCTGCTGAAGCCAGTTACC-TTCGGAAAAAGAGTTGGTAGCTCTTTGATNCGGCAACAA 546
|||||
QY 1227 ACCACCGCTGGTAGCGTGGTGTGTTTGTTCGACGACGACAGATT-ACGCGCAGAAAAA 1285
|||||
Db 547 ACCACCGCTGGTAGC-GTGGTTTTTTTGTTCGACGACGACAGATTAAACGCGNGGATTCA 605
|||||
QY 1286 AGGATC 1291
|||||
Db 606 AGGATC 611
|||||

RESULT 15
AL044364 AL044364 954 bp mRNA EST 29-FEB-2000
LOCUS
DEFINITION DKFZp434C172.s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZp434C172.3', mRNA sequence.
VERSION AL044364
KEYWORDS EST.
SOURCE AL044364.1 GI:5432586
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS Ansorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Ansorge, Benes, et al.)
Unpublished (1999)
CONTACT: Ansorge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ): Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories, of the
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN
```


Query Match	35.6%;	Score 550.2;	DB 10;	Length 954;
Best Local Similarity	99.5%;	Pred. No. 3.5e-147;		
Matches 552;	Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
QY	690	GCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA	749	
Db	400	GCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA	459	
QY	750	CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACGAGCGTTTCCCCCT	809	
Db	460	CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACGAGCGTTTCCCCCT	519	
QY	810	GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCGCGTTACCGGATACCTGTCCGCC	869	
Db	520	GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCGCGTTACCGGATACCTGTCCGCC	579	
QY	870	TTTCTCCCTTCGGGAGCGTGGCGCTTTCATATGCTACGCTGTAGGTATCTCAGTTTCG	929	
Db	580	TTTCTCCCTTCGGGAGCGTGGCGCTTTCATATGCTACGCTGTAGGTATCTCAGTTTCG	639	
QY	930	GTGTAGGTCGTTCCGCTCCAGCTGGGCTGTGTGCAGCAACCCCGCTTACGCCGACCGC	989	
Db	640	GTGTAGGTCGTTCCGCTCCAGCTGGGCTGTGTGCAGCAACCCCGCTTACGCCGACCGC	699	
QY	990	TGCGCCTTATCCGGTAACCTATCTGTCTTGTCTCAACCCGGTAAGACACGACTTATCGCCA	1049	
Db	700	TGCGCCTTATCCGGTAACCTATCTGTCTTGTCTCAACCCGGTAAGACACGACTTATCGCCA	759	
QY	1050	CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG	1109	
Db	760	CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG	819	
QY	1110	TTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGTATTTGGTATCTCGGCT	1169	
Db	820	TTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGTATTTGGTATCTCGGCT	879	
QY	1170	CTGCTGAAGCCAGTTACCTTTCGAAAAAGATTGGTAGCTCTTGTATCCGGCAACAACC	1229	
Db	880	CTGCTGAAGCCAGTTACCTTTCGAAAAAGATTGGTAGCTCTTGTATCCGGCAACAACC	939	
QY	1230	ACCGCTGGTAGCGGT	1244	
Db	940	ACCGCTGGTAGCGGT	954	

Search completed: January 17, 2002, 08:54:31
Job time: 7143 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2002, 11:50:28 ; Search time 10436.3 Seconds
(without alignments)
2856.404 Million cell updates/sec

Title: US-09-242-202a-28
Perfect score: 1807
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTCCACAGCCTCTCCACACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo_hum.*
- 31: em.htgo_inv.*
- 32: em.htgo_rod.*
- 33: em.htg_hum.*
- 34: em.htg_inv.*
- 35: em.htg_rod.*
- 36: em.htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	765	42.3	885	12	SYNPIAN7V	L08875 PIAN7 cloni
2	756.2	41.8	902	12	ATPIVX	V00089 Cloning vec
3	756.2	41.8	902	12	PRIVX	X14353 Plasmid pi-
4	756.2	41.8	902	12	SYNPIVXV	L08918 PI-VX cloni
5	755	41.8	2932	6	I07209	I07209 Sequence 3
6	656.2	36.3	3392	12	AB008864	AB008864 Expressio
7	607.4	33.6	7252	6	AX003206	AX003206 Sequence
8	604.4	33.4	5865	6	A60212	A60212 Sequence 8
9	604.4	33.4	5865	6	AR122288	AR122288 Sequence
10	604.4	33.4	6028	6	A60209	A60209 Sequence 5
11	604.4	33.4	6028	6	AR122285	AR122285 Sequence
12	604.4	33.4	6061	6	A60210	A60210 Sequence 6
13	604.4	33.4	6061	6	AR122286	AR122286 Sequence
14	604.4	33.4	6312	6	A60211	A60211 Sequence 7
15	604.4	33.4	6312	6	AR122287	AR122287 Sequence
16	604.2	33.4	11795	6	AX027785	AX027785 Sequence
17	604.2	33.4	13254	6	AR038307	AR038307 Sequence
18	604.2	33.4	13254	6	AR038321	AR038321 Sequence
19	604.2	33.4	13254	6	I58596	I58596 Sequence 15
20	604.2	33.4	13254	6	I58610	I58610 Sequence 17
21	603	33.4	1905	6	AR027070	AR027070 Sequence
22	603	33.4	1905	6	I86203	I86203 Sequence 9
23	603	33.4	1968	12	AF087567	AF087567 Cloning v
24	603	33.4	1969	12	AF086840	AF086840 Cloning v
25	603	33.4	1969	12	AF086841	AF086841 Cloning v
26	603	33.4	1969	12	AF086842	AF086842 Cloning v
27	603	33.4	1969	12	AF086843	AF086843 Cloning v
28	603	33.4	1969	12	AF086844	AF086844 Cloning v
29	603	33.4	1969	12	AF086845	AF086845 Cloning v
30	603	33.4	1969	12	AF086846	AF086846 Cloning v
31	603	33.4	1969	12	AF086847	AF086847 Cloning v
32	603	33.4	1969	12	AF086848	AF086848 Cloning v
33	603	33.4	1969	12	AF086849	AF086849 Cloning v
34	603	33.4	1969	12	AF086850	AF086850 Cloning v
35	603	33.4	1969	12	AF086851	AF086851 Cloning v
36	603	33.4	1969	12	AF086852	AF086852 Cloning v
37	603	33.4	1969	12	AF086853	AF086853 Cloning v
38	603	33.4	1969	12	AF087562	AF087562 Cloning v
39	603	33.4	1969	12	AF087563	AF087563 Cloning v
40	603	33.4	1969	12	AF087564	AF087564 Cloning v
41	603	33.4	1969	12	AF087565	AF087565 Cloning v
42	603	33.4	1969	12	AF087566	AF087566 Cloning v
43	603	33.4	1969	12	AF087564	AF087564 Cloning v
44	603	33.4	1969	12	AF087565	AF087565 Cloning v
45	603	33.4	1969	12	AF087566	AF087566 Cloning v

ALIGNMENTS

RESULT 1

SYNPIAN7V

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

SYNPIAN7V 885 bp
PIAN7 cloning vector.
L08875
L08875.1 GI:310776
Synthetic construct DNA.
synthetic construct
artificial sequence.
1 (bases 1 to 885)
Gilbert,W.
Obtained from VecBase 3.0
Unpublished (1991)
These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program. PIAN7 -
Cloning vector
ENTRY PIAN7
#TYPE DNA CIRCULAR
TITLE PIAN7 - Cloning vector
DATE 12-SEP-1986

```

#sequence 16-DEC-1986
ACCESSION VB0066
SOURCE artificial
REFERENCE
#number 1
#citation sequence information from New England Biolabs COMMENT
Obtained 12-SEP-1986 from New England Biolabs
by magnetic tape
Revised 16-DEC-1986 by F. Pfeiffer:
449/450 'AT' to 'TA' to match revised sequence of PBR322 COMMENT
PIA7 is thought to replace pIVX.
COMMENT
The polylinker of PIA7 contains additional BglII and XbaI
sites within the M13mp8/pUC8 polylinker.
KEYWORDS
CROSSREFERENCE
#parent
VecBase(3):pIVX
#parent
VecBase(3):PBR322, VecSource(3):Poly8, GenBank(50):Ecofgy
PARENT
Features of PIA7 (885 bp)
residue source
1-202 202-1 (c) Tyr-tRNA synthetic (GenBank(50):Ecofgy)
198-225 1-28 part 1 of pUC8/M13mp8-polylinker
234-239 31-36 part 2 of pUC8/M13mp8-polylinker
240-880 2522-3162 pBR322
Conflict (cfl) and Mutations (mut): none
FEATURE
POLYLINKER EcoRI-SmaI-BamHI-SalI-PstI-BglII-XbaI-HindIII SELECTION
#suppressor Supf
SUMMARY PIA7 #length 885 #checksum 9656.
Location/Qualifiers
1..885
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 193 a 239 c 236 g 217 t
ORIGIN
Query Match 42.3%; Score 765; DB 12; Length 885;
Best Local Similarity 95.9%; Pred. No. 2.6e-211;
Matches 805; Conservative 0; Mismatches 10; Indels 24; Gaps 1;
QY 478 TTTCGGACTTTTGAAGTGTGTTGGGGGAAGATTGCAACCTTCGAAGTCGATGAC 537
DB 1 TTTCGGACTTTTGAAGTGTGTTGGGGGAAGATTGCAACCTTCGAAGTCGATGAC 60
QY 538 GCAGATTAGAGTCTGCTCCCTTGGCGCTCGGGAACCCACACCGGTAATGCTTTT 597
DB 61 GCAGATTAGAGTCTGCTCCCTTGGCGCTCGGGAACCCACACCGGTAATGCTTTT 120
QY 598 ACTGCGCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAATCAGCGCGCTGTAA 657
DB 121 ACTGCGCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAATCAGCGCGCTGTAA 180
QY 658 AGTGTTACGTTGAGAAGATTTC-----CTGACGCCGCCG 693
DB 181 AGTGTTACGTTGAGAAGATTTC-----CTGACCGTGCAGATCTCTAGAAGCTTG 240
QY 694 CGTGTGCGGCTTTTCCATAGCTCGCCCTTACGAGCATCACAAAATCGACGCT 753
DB 241 CGTGTGCGGCTTTTCCATAGCTCGCCCTTACGAGCATCACAAAATCGACGCT 300
QY 754 CAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCTCGAA 813
DB 301 CAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCTCGAA 360
QY 814 GCTCCCTCGTGCCTCTCCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTC 873
DB 361 GCTCCCTCGTGCCTCTCCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTC 420
QY 874 TCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTCTAGTATCTCAGTTCGCTGT 933

```

```

Db 421 TCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCGGTGT 480
QY 934 AGGTCGTTCCGCTCCAAAGCTGGCTGTGTGCAGCAACCCCGCTTTCAGCCCGACGCTCG 993
DB 481 AGGTCGTTCCGCTCCAAAGCTGGCTGTGTGCAGCAACCCCGCTTTCAGCCCGACGCTCG 540
QY 994 CCTATCCGGTAACTATATCTTGTAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGG 1053
DB 541 CCTATCCGGTAACTATATCTTGTAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGG 600
QY 1054 CAGCAGCCACTGGTAAACAGGATTACGAGCGAGGTATGTAGGCGGTCTACAGAGTTCT 1113
DB 601 CAGCAGCCACTGGTAAACAGGATTACGAGCGAGGTATGTAGGCGGTCTACAGAGTTCT 660
QY 1114 TGAAGTGTGCTGCTAACTACGCTACACTAGAAAGACAGTATTTGGTATCTCGCGTCTGC 1173
DB 661 TGAAGTGTGCTGCTAACTACGCTACACTAGAAAGACAGTATTTGGTATCTCGCGTCTGC 720
QY 1174 TGAAGCCAGTTACCTTCGGAAAAGAGTTGTAGTCTTTGATCCGCGCAACAAACACCG 1233
DB 721 TGAAGCCAGTTACCTTCGGAAAAGAGTTGTAGTCTTTGATCCGCGCAACAAACACCG 780
QY 1234 CTGGTAGCGGTGTTTTTTTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGATCT 1292
DB 781 CTGGTAGCGGTGTTTTTTTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGATCT 839
RESULT 2
ATPIVX 902 bp DNA SYN 28-SEP-1992
LOCUS Cloning vector pi-VX used for screening bacteriophage lambda gene
DEFINITION libraries for specific DNA sequences in Escherichia coli. Probe
sequences are inserted in the vector pi-VX. Contains a polylinker,
and origin of replication (derived from pMB1) and a tyrosine
amber-suppressor gene (synthetic supF gene).
ACCESSION V00089
VERSION V00089.1 GI:58159
KEYWORDS cloning vector; origin of replication; plasmid; polylinker;
suppressor gene.
SOURCE Cloning vector pi-VX.
ORGANISM Cloning vector pi-VX
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 902)
AUTHORS Maniatis,T., Fritsch,E.F. and Sambrook,J.
TITLE Molecular cloning
JOURNAL (in) Maniatis,T., Fritsch,E.F. and Sambrook,J. (Eds.);
MOLECULAR CLONING: 353-361;
Cold Spring Harbor Laboratory (1982)
REFERENCE 2 (bases 1 to 902)
AUTHORS Fitzwater,T., Yang,X.L., Zhang,X.Y. and Polisky,B.
TITLE Mutations affecting RNA-DNA hybrid formation of the Cole1
replication primer RNA. Restoration of RNA I sensitivity to a
copy-number mutant by second-site mutations
JOURNAL J. Mol. Biol. 226 (4), 997-1008 (1992)
MEDLINE 92389346
FEATURES Location/Qualifiers
source
1..902
/organism="Cloning vector pi-VX"
/db_xref="taxon:31856"
rep_origin 600
/note="origin of replication"
BASE COUNT 192 a 250 c 237 g 223 t
ORIGIN
Query Match 41.8%; Score 756.2; DB 12; Length 902;
Best Local Similarity 98.1%; Pred. No. 9.5e-209;
Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
QY 471 CGAATTCCTTCGACCTTTTGAAGTGTGTTGGGGGAAGGATTCGAACTTCGAAGT 530
DB 109 CGAATTCCTTCGACCTTTTGAAGTGTGTTGGGGGAAGGATTCGAACTTCGAAGT 168

```

```

QY 531 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA 590
DB 169 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA 228
QY 591 TGTCTTTTACTGGCGCTGCTCCCTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC 650
DB 229 TGTCTTTTACTGGCGCTGCTCCCTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC 288
QY 651 GCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGGTTTTTC 710
DB 289 GCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGGTTTTTC 337
QY 711 CATAGGCTCGCCCGCTGACGAGATCAAAAATGACGCTCAAGTCAGAGGTGCGGA 770
DB 338 CATAGGCTCGCCCGCTGACGAGATCAAAAATGACGCTCAAGTCAGAGGTGCGGA 397
QY 771 AACCCGACAGGACTATAAAGATACGAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT 830
DB 398 AACCCGACAGGACTATAAAGATACGAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT 457
QY 831 CTTGTTCCGACCTCGCGCTTACGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 890
DB 458 CTTGTTCCGACCTCGCGCTTACGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 517
QY 891 GCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGGTGAGGTGCTTCCGTCCTCAAG 950
DB 518 GCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGGTGAGGTGCTTCCGTCCTCAAG 577
QY 951 CTGGGCTGTGTGACGAACCCCGCTTACGCGGACGCTGTCGCCCTTATCCGGTAACTAT 1010
DB 578 CTGGGCTGTGTGACGAACCCCGCTTACGCGGACGCTGTCGCCCTTATCCGGTAACTAT 637
QY 1011 CGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAA 1070
DB 638 CGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAA 697
QY 1071 AGGATTAGCAGGAGGATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTTAAC 1130
DB 698 AGGATTAGCAGGAGGATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTTAAC 757
QY 1131 TAGGCTTACACTAGAAGCAGATTTGGTATCTGCTGCGCTGCTGCGTGAAGCAGTTACCTTC 1190
DB 758 TAGGCTTACACTAGAAGCAGATTTGGTATCTGCTGCGCTGCTGCGTGAAGCAGTT-CCTTC 816
QY 1191 GGAAGAGAGTGTGAGTCTTGTATCGGCAACAAACACCCGCTGCTAGCGGTGTTTT 1250
DB 817 GGAAGAGAGTGTGAGTCTTGTATCGGCAACAAACACCCGCTGCTAGCGGTGTTTT 876
QY 1251 TTTGTTTGAAGCAGCAGATTACGC 1275
DB 877 TTTGTTTGAAGCAGCAGATTACGC 901

RESULT 3
LOCUS PPIVX
DEFINITION PPIVX 902 bp DNA SYN 07-FEB-1990
ACCESSION X14353
VERSION X14353.1 GI:58268
KEYWORDS artificial sequence; origin of replication; plasmid; polylinker;
suppressor transfer RNA.
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 902)
AUTHORS Seed,B.
TITLE Purification of genomic sequences from bacteriophage libraries by
recombination and selection in vivo
JOURNAL Nucleic Acids Res. 11 (8), 2427-2445 (1983)
MEDLINE 83220795
FEATURES
source
location
1..902

```

```

/organism="synthetic construct"
/plasmid="pi-VX"
/db_xref="taxon:32630"
misc_feature 1..109
trna 110..212
misc_feature 213..902
BASE COUNT 192 a 249 c 238 g 223 t
ORIGIN
Query Match 41.8%; Score 756.2; DB 12; Length 902;
Best Local Similarity 98.1%; Pred. No. 9.5e-209;
Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
QY 471 CGAATCTTTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAAGT 530
DB 109 CGAATCTTTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAAGT 168
QY 531 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA 590
DB 169 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA 228
QY 591 TGTCTTTTACTGGCGCTGCTCCCTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC 650
DB 229 TGTCTTTTACTGGCGCTGCTCCCTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC 288
QY 651 GCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGGTTTTTC 710
DB 289 GCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGGTTTTTC 337
QY 711 CATAGGCTCGCCCGCTGACGAGATCAAAAATGACGCTCAAGTCAGAGGTGCGGA 770
DB 338 CATAGGCTCGCCCGCTGACGAGATCAAAAATGACGCTCAAGTCAGAGGTGCGGA 397
QY 771 AACCCGACAGGACTATAAAGATACGAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT 457
DB 398 AACCCGACAGGACTATAAAGATACGAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT 457
QY 831 CTTGTTCCGACCTCGCGCTTACGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 890
DB 458 CTTGTTCCGACCTCGCGCTTACGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 517
QY 891 GCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGGTGAGGTGCTTCCGTCCTCAAG 950
DB 518 GCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGGTGAGGTGCTTCCGTCCTCAAG 577
QY 951 CTGGGCTGTGTGACGAACCCCGCTTACGCGGACGCTGTCGCCCTTATCCGGTAACTAT 1010
DB 578 CTGGGCTGTGTGACGAACCCCGCTTACGCGGACGCTGTCGCCCTTATCCGGTAACTAT 637
QY 1011 CGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAA 1070
DB 638 CGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAA 697
QY 1071 AGGATTAGCAGGAGGATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTTAAC 1130
DB 698 AGGATTAGCAGGAGGATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTTAAC 757
QY 1131 TAGGCTTACACTAGAAGCAGATTTGGTATCTGCTGCGCTGCTGCGTGAAGCAGTTACCTTC 1190
DB 758 TAGGCTTACACTAGAAGCAGATTTGGTATCTGCTGCGCTGCTGCGTGAAGCAGTT-CCTTC 816
QY 1191 GGAAGAGAGTGTGAGTCTTGTATCGGCAACAAACACCCGCTGCTAGCGGTGTTTT 1250
DB 817 GGAAGAGAGTGTGAGTCTTGTATCGGCAACAAACACCCGCTGCTAGCGGTGTTTT 876
QY 1251 TTTGTTTGAAGCAGCAGATTACGC 1275
DB 877 TTTGTTTGAAGCAGCAGATTACGC 901

```

RESULT 4

SYNPIVXV 902 bp DNA circular SYN 26-JUL-1993
 PI-VX cloning vector used for screening bacteriophage lambda.
 L08918
 L08918.1 GI:310786
 Synthetic construct DNA.
 synthetic construct
 artificial sequence.
 1 (bases 1 to 902)
 Gilbert, W.
 Obtained from VecBase 3.0
 Unpublished (1991)
 These data and their annotation were supplied to GenBank by Will
 Gilbert under the auspices of the GenBank Curator Program. PI-VX -
 Cloning vector used for screening bacteriophage lambda ENTRY PIVX
 Cloning vector used for screening bacteriophage
 #TYPE DNA CIRCULAR
 TITLE PI-VX - Cloning vector used for screening bacteriophage
 lambda
 DATE 17-DEC-1982
 ACCESSION V00010
 SOURCE artificial
 COLLECTION ATCC 39083
 REFERENCE
 #number 1
 #authors Seed B.
 #journal Nucl. Acids Res. (1983) 11: 2427-2445
 #title
 Purification of genomic sequences from bacteriophage libraries
 and selection in vivo.
 #comment bases 1-902
 #number 2
 #authors Maniatis T., Fritsch E.F., Sambrook J.
 #book (IN) Maniatis T., Fritsch E.F., Sambrook J.;
 MOLECULAR CLONING: 353-361;
 COLD SPRING HARBOR LABORATORY (1982).
 #comment BASES 1 TO 902
 REFERENCE 3
 #authors Backman K., Betlach M., Boyer H.W., Yanofsky S.
 #journal Cold Spring Harb Symp Quant Biol 43, 69-76 (1980)
 #title
 Genetic and Physical Studies on the replication of Cole1-type
 plasmids
 #comment
 the sequence of the PIVX origin is from pBR345 as published in
 this paper
 COMMENT
 The origin of replication is derived from pMB1 via pBR345 and
 pK413. The
 sequence given for pBR345 differs at one position from that of
 pBR322. This
 may be a sequencing error.
 COMMENT
 The ATCC strain 39083 contains PIVX in combination with plasmid
 p3.
 COMMENT
 from GenBank
 USED FOR SCREENING BACTERIOPHAGE LAMBDA GENE LIBRARIES FOR
 SPECIFIC DNA SEQUENCES IN ESCHERICHIA COLI. PROBE SEQUENCES
 ARE INSERTED IN THE VECTOR PI-VX. CONTAINS A POLYLINKER, AND
 ORIGIN OF REPLICATION AND A TYROSINE
 AMBER-SUPPRESSOR GENE (SYNTHETIC SUPP GENE).
 KEY FROM TO
 ORGRPL 600 600 ORIGIN OF REPLICATION KEYWORDS
 suppressor gene \
 polylinker \
 origin of replication
 CROSSREFERENCE
 #identical
 GenBank(50):PIVX, EMBL(11):ATPIVX
 #parent

VecBase(3):pBR322, GenBank(50):EcoTgy, VecSource(3):OriMB1
 #offspring
 VecBase(3):PIAN7, VecBase(3):mWB2344
 PARENT
 Features pf PIVX (902 bp)
 residue source
 4- 37 1- 34 pBR322
 1- 115 1- 115 PIVX-Polylinker
 111- 317 207- 1 (c) Tvr-trNA synthetic (GenBank(50):EcoTgy)
 320- 901 1- 583 OriMB1
 320- 901 2522-3104 pBR322
 Conflict (cfl) and Mutations (mut):
 PIVX source
 cfl 811-12 T.C TAC 492-94 OriMB1
 cfl 811-12 T.C TAC 3013-15 pBR322
 FEATURE
 SUPP
 POLYLINKER EcoRI-ClaI-HindIII-XbaI-BglII-PstI-BamHI-EcoRI
 SELECTION
 #suppressor SupF
 SUMMARY PIVX #length 902 #checksum 9265.
 Location/Qualifiers
 1..902
 /organism="synthetic construct"
 /db_xref="taxon:32630" 223 t
 BASE COUNT 192 a 250 c 237 g
 ORIGIN

Query Match 41.8%; Score 756.2; DB 12; Length 902;
 Best Local Similarity 98.1%; Pred. No. 9.5e-209;
 Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
 QY 471 CGAATCTTTTCGGACCTTTGAAAGTGATGGTGGGGGAAGGATTTCGAACCTTCGAAGT 530
 Db 109 CGAATCTTTTCGGACCTTTGAAAGTGATGGTGGGGGAAGGATTTCGAACCTTCGAAGT 168
 QY 531 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAAACCCACCACGCGTAA 590
 Db 169 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAAACCCACCACGCGTAA 228
 QY 591 TCGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGCC 650
 Db 229 TCGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGCC 288
 QY 651 GCTGTAAGTGTGTACGTTGAGAAAGATTCTCTGCAGCCCGCGCTTCTGCGGCTTTTTC 710
 Db 289 GCTGTAAGTGTGTACGTTGAGAAAGATTCTCTGCAGCCCGCGCTTCTGCGGCTTTTTC 337
 QY 711 CATAGGCTCGGCCCTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGA 770
 Db 338 CATAGGCTCGGCCCTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGA 397
 QY 771 AACCCGACAGGACTATAAGATACACAGGCGTTTCCCGCTTGGAAAGCTCCCTCGCGCTCT 830
 Db 398 AACCCGACAGGACTATAAGATACACAGGCGTTTCCCGCTTGGAAAGCTCCCTCGCGCTCT 457
 QY 831 CTGTTCGACCGCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 890
 Db 458 CTGTTCGACCGCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 517
 QY 891 GCCTTTCTCAATGCTCACGCTGTAGTATCTCAGTTTCGGGTGTAGGTCTGCTCCTCAAG 950
 Db 518 GCCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGGTGTAGGTCTGCTCCTCAAG 577
 QY 951 CTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTTACGCCCTTATCCGGTAACTAT 1010
 Db 578 CTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTTACGCCCTTATCCGGTAACTAT 637
 QY 1011 CGCTTTGAGTCCAAACCGGTAAGACAGCACTTATCGCACTTGGCAGCAGCACTGGTAAC 1070
 Db 638 CGCTTTGAGTCCAAACCGGTAAGACAGCACTTATCGCACTTGGCAGCAGCACTGGTAAC 697

```
Qy 1071 AGGATTAGCAGAGCAGGATATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAAC 1130
|||||
Db 698 AGGATTAGCAGAGCAGGATATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAAC 757
|||||
Qy 1131 TACGGCTACACTAGAAGCAGATATTTGGTATCTGCCCTCTGCTGAAGCAGATTACCTTC 1190
|||||
Db 758 TACGGCTACACTAGAAGCAGATATTTGGTATCTGCCCTCTGCTGAAGCAGATT-CCTTC 816
|||||
Qy 1191 GGAAGAGAGTTGGTACCTCTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTTTTT 1250
|||||
Db 817 GGAAGAGAGTTGGTACCTCTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTTTTT 876
|||||
Qy 1251 TTGTTTGCAGCAGCAGATTACGC 1275
|||||
Db 877 TTTGTTTGCAGCAGCAGATTACGC 901
|||||

RESULT 5
107209/c 2932 bp PAT 02-DEC-1994
LOCUS Sequence 3 from Patent EP 0330191.
DEFINITION
ACCESSION 107209
VERSION 107209.1 GI:590060
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2932)
AUTHORS Seed,B.D.,D.O.M.B., Allen,J., Aruffo,A., Camerini,D., Lauffer,L.D.,
Oquendo,C.P., Simmons,D., Stamenkovic,I. and Stengelin,S.D.
TITLE Rapid immunoselection cloning method
JOURNAL Patent: EP 0330191-A2 3 30-AUG-1989;
FEATURES
source Location/Qualifiers
BASE COUNT 774 a 699 c 682 g 777 t
ORIGIN

Query Match 41.88; Score 755; DB 6; Length 2932;
Best Local Similarity 97.68; Pred. No. 2.2e+108;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

Qy 469 ATCGAATTTCTTCGACCTTTTGAAGTGATGGTGGGGGAAGGATTGCAACCTTCGAA 528
|||||
Db 804 AGCAGATCTTTCGACCTTTTGAAGTGATGGTGGGGGAAGGATTGCAACCTTCGAA 745
|||||
Qy 529 GTCGATGACGCGCAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCCACCGGT 588
|||||
Db 744 GTCGATGACGCGCAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCCACCGGT 685
|||||
Qy 589 AATGCTTTTACTGCGCTGCTCCCTTATCGGAAGCGGGGGGCATCATATCAATGACGCG 648
|||||
Db 684 AATGCTTTTACTGCGCTGCTCCCTTATCGGAAGCGGGGGGCATCATATCAATGACGCG 625
|||||
Qy 649 CCGCTGTAAGTGTTAGTTGAGAAAGTAATTCCTGCAGCCCGCGCTGCTGCGCTTTT 708
|||||
Db 624 CCGCTGTAAGTGTTAGTTGAGAAAGA---CCGCGTAATTCGCGCTGCTGCGCTTTT 568
|||||
Qy 709 TCCATAGGCTCCGCGCCCTCGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGC 768
|||||
Db 567 TCCATAGGCTCCGCGCCCTCGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGC 508
|||||
Qy 769 GAAACCGCAGGAGCTATAAGATACACAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCT 828
|||||
Db 507 GAAACCGCAGGAGCTATAAGATACACAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCT 448
|||||
Qy 829 CTCCTGTTCCGACCTCCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAAGCG 888
|||||
Db 447 CTCCTGTTCCGACCTCCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAAGCG 388
|||||
Qy 889 TGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCGGTAGTCTGCTTCGCTCCA 948
|||||
```

```
Db 387 TGGCGCTTTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCCGTGTAGGTGCTTCGCTCCA 328
Qy 949 AGCTGGGCTGTGTGCAGAAACCCCGCTTCAGCCCGCAGCGCTTATTCGGGTAAC 1008
|||||
Db 327 AGCTGGGCTGTGTGCAGAAACCCCGCTTCAGCCCGCAGCGCTTATTCGGGTAAC 268
|||||
Qy 1009 ATCGCTTTGAGTCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTA 1068
|||||
Db 267 ATCGCTTTGAGTCCAAACCCCGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTA 208
|||||
Qy 1069 ACAGGATTAGCAGAGCAGGATATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 1128
|||||
Db 207 ACAGGATTAGCAGAGCAGGATATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 148
|||||
Qy 1129 ACTACGCTACACTAGAAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCAGCTTACT 1188
|||||
Db 147 ACTACGCTACACTAGAAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCAGCTT-CCT 89
|||||
Qy 1189 TCGGAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTT 1248
|||||
Db 88 TCGGAAAAGAGTTGGTAGCTTCTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTT 29
|||||
Qy 1249 TTTTGTGTTGCAACGACAGATTACGC 1275
|||||
Db 28 TTTTGTGTTGCAACGACAGATTACGC 2
|||||

RESULT 6
AB009864
LOCUS 3392 bp DNA circular SYN 26-DEC-1997
DEFINITION Expression vector pME18S-FL3, complete sequence.
ACCESSION AB009864
VERSION AB009864.1 GI:2723416
KEYWORDS
SOURCE expression vectors plasmid:pME18S-FL3 DNA.
ORGANISM Plasmid vectors
REFERENCE 1 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, The Institute of Medical Science, University of Tokyo,
Department of Virology; 4-6-1, Shirokanedai, Minato-ku, Tokyo 108,
Japan (E-mail: ssugano@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
REFERENCE 2 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE pME18S-FL3: a versatile expression vector
JOURNAL Published Only in Database (1997) In press
FEATURES
source Location/Qualifiers
1..3392 /organism="vectors"
/promoter /plasmid="pME18S-FL3"
/polyA_site /db_xref="taxon:29278"
BASE COUNT 810 a 902 c 818 g 862 t
ORIGIN

Query Match 36.38; Score 656.2; DB 12; Length 3392;
Best Local Similarity 70.88; Pred. No. 1.1e-179;
Matches 990; Conservative 0; Mismatches 288; Indels 121; Gaps 4;

Qy 15 TGGCGCGGATTTCTTATCAGTGAAGTTGGTGGACATATTATGTTTATCAGTGATAAG 74
|||||
Db 828 TGGCGCGGATTTCTTATCAGTGAAGTTGGTGGACATATTATGTTTATCAGTGATAAG 887
|||||
Qy 75 TGTCAGCATGACAAAGTTGCAGCCGAATACAGTATCCGTGCGCGCCCTGGAGCTGTGA 134
|||||
Db 888 TGTCAGCATGACAAAGTTGCAGCCGAATACAGTATCCGTGCGCGCCCTAGACCTGTGA 947
|||||
Qy 135 ACGAGGTGCGGTAGACGGTCTGACGACACCGCAAACTGCGCGGAACGGTTGGGGGTGCAGC 194
|||||
```

Db 948 ACAGGTCGCGTAGACGGTCTACGACACGCAAACTGGCGGAACGGTTGGGGTTTCAGC 1007
QY 195 AGCGCGCGCTTTACTGCGACTTCAGGAACAAGCGGGCGCTTAAAGGGCCATATGGTGAGT 254
Db 1008 AGCGCGCGCTTTACTGCGACTTCAGGAACAAGCGGGCGCTGCTCGACGCACTGGCCGAAG 1067
QY 255 GGATGC-----CATTGACCCAGCGGGGATGGGGGAGACC 289
Db 1068 CCATGCTGGCGGAGAAATCATAGACATTCGGTGGCGAGAGCGGACGACTGGCGTCAAT 1127
QY 290 TGTAGTCAGAGCCCCCGGCGAGCAGACGCAATGCCGCTCTTCCCTGCGAGATGAGTA 349
Db 1128 TTCTGACTGGGAATGCCCGCAGCTTCAGGCAGCGCTGCTCGCCTACCGCCACACAATG 1187
QY 350 GTGAGTCCCTCTCGTGGCCCTGGAAGTTGCCACTCCAGTCCAGTCCCAACGAGCTTGTCT-- 406
Db 1188 GCACCCCGTCTCGAGCTGCGAGTCGCGGCGCTGAGCTAGTCTAGAGAAAAACCTCCC 1247
QY 407 -----AATAAAATTAAGTTGCATCATTTTCTCTGACTAGTGT 444
Db 1248 ACACCTCCCTGNACTGAAACATAAAATGAATGCAATTTGTTGTTAACTTGTTAT 1307
QY 445 CCTCTAATATATTAAGTTGATATGAAATCTTTTGGGACTTTTGAAGTATGATGGTGT 504
Db 1308 TGCAGCTTATAATGTTTACAAATAAAGCAATAGCATCACAAATTTCAAAATAAAGCAAT 1367
QY 505 GGGGAGGATTCGAACCTTCGAAGTCGATGAGCGGAGATTTAGAGTCTGCTCCCTTTGG 564
Db 1368 TTTTTCACGATTCATGATGTTGTGTTGTCCAAACTCATCAATGATCTTATCATGTCGT 1427
QY 565 CCGCTCGGGAACCCACGCGGTAATGCTTTTACTGGCTGTCTCCCTTATCGGGAAGCG 624
Db 1428 GATCCCGGGTACCGAGCTCGAATTAATCTCTTCGCTCTCTCGCTCACTGACTGCT 1487
QY 625 GGGCGCATCATATAAANAGCGG----- 648
Db 1488 GCGCTCGGTGCTTCGGCTGGCGGCGAGCGGTATCAGCTCACTCAAGGCGGTAATACGTT 1547
QY 649 -----CCGCTGTAAAGTGTACGTTGAGAAACAATTCCTGACGCC 688
Db 1548 ATCCACAGAAATCAGGGATACGCGAGGAAGAACATGTGAGCAAAAGCGCAGCAAAAGC 1607
QY 689 C-----GCCGGTTGCTGGCGTTTTCATAGGCTCGCCCCCTGACGA 733
Db 1608 CAGGAACGTAATAAAGCGGTTGCTGGCGTTTTCATAGGCTCGCCCCCTGACGA 1667
QY 734 GCATCAAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATAAAGATA 793
Db 1668 GCATCAAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATAAAGATA 1727
QY 794 CCAGCGTTTCCCTCGAAGCTCCCTGTCGCTCTCTGTTCCGACCTCGCGCTTAC 853
Db 1728 CCAGCGTTTCCCTCGAAGCTCCCTGTCGCTCTCTGTTCCGACCTCGCGCTTAC 1787
QY 854 CGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGGTTTCTCAATGCTCACGCTG 913
Db 1788 CGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGGTTTCTCAATGCTCACGCTG 1847
QY 914 TAGGTATCTCAGTTCGGTGTAGTGTTCGCTCAAGTGGGCTGTGTGACGAACCCCC 973
Db 1848 TAGGTATCTCAGTTCGGTGTAGTGTTCGCTCAAGTGGGCTGTGTGACGAACCCCC 1907
QY 974 GTTCAGCCCGACCGCTGCGCTTATCCGGTAACATATCTGCTTCAGTCCAAACCGGTAAG 1033
Db 1908 GTTCAGCCCGACCGCTGCGCTTATCCGGTAACATATCTGCTTCAGTCCAAACCGGTAAG 1967
QY 1034 ACAGGACTTATCGCACTGGCAGAGCCACTTGTAAACAGGATTAAGCAGCAGGATATGT 1093
Db 1968 ACAGGACTTATCGCACTGGCAGAGCCACTTGTAAACAGGATTAAGCAGCAGGATATGT 2027
QY 1094 AGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAAGACGT 1153

Db 2028 AGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGT 2087
QY 1154 ATTGTGTAATCTGCTGCTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTG 1213
Db 2088 ATTGTGTAATCTGCTGCTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTG 2147
QY 1214 ATCCGCAAAACAAACACCGCTGCTGAGCGGTGTTTGTGTTTGAAGCAGCAGATTAC 1273
Db 2148 ATCCGCAAAACAAACACCGCTGCTGAGCGGTGTTTGTGTTTGAAGCAGCAGATTAC 2207
QY 1274 GCGCAGAAAAAAGGATCT 1292
Db 2208 GCGCAGAAAAAAGGATCT 2226
RESULT 7
AX003206 AX003206 7252 bp DNA PAT 24-AUG-2000
LOCUS Sequence 57 from Patent WO9932646.
DEFINITION AX003206
ACCESSION AX003206
VERSION AX003206.1 GI:9927068
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 7252)
AUTHORS Carroll,M.W. and Mitrophanous,K.
TITLE Equine infectious anaemia virus (eia) based
JOURNAL Patent: WO 9932646-A 57 01-JUL-1999;
CARROLL MILES WILLIAM (GB); MITROPHANOUS KYRIACOS (GB)
FEATURES
Location/Qualifiers
1..7252
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="complete sequence of pSC65"
BASE COUNT 1942 a 1686 c 1751 g 1873 t
ORIGIN
Query Match 33.6%; Score 607.4; DB 6; Length 7252;
Best Local Similarity 97.5%; Pred. No. 2.1e-165;
Matches 617; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 660 TGTTCAGTTGAGAAAGAAATTCCTGCGAGCCGCGCGGTGCTGCGGTTTTCATAGGCTC 719
Db 5062 TATACCGGTGAGAACATGAGACTGACTCGGGCGCGGTGCTGCGGTTTTCATAGGCTC 5121
QY 720 CGCCCGCTCACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 779
Db 5122 CGCCCGCTCACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 5181
QY 780 GGACTATAAAGATACAGGCGTTTCCCGCTTGAAGCTCCCTGCTGCGCTCTCCTGTTCCG 839
Db 5182 GGACTATAAAGATACAGGCGTTTCCCGCTTGAAGCTCCCGCTGCGCTCTCCTGTTCCG 5241
QY 840 ACCCTGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGTTTCT 899
Db 5242 ACCCTGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGTTTCT 5301
QY 900 CAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCAAAGCTGGGCTGT 959
Db 5302 CAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCAAAGCTGGGCTGT 5361
QY 960 GTGCACGAACCCCGCTTCAGCCCGACCGCTGCGGCTTATCCGGTAACATATCGCTTTGAG 1019
Db 5362 GTGCACGAACCCCGCTTCAGCCCGACCGCTGCGGCTTATCCGGTAACATATCGCTTTGAG 5421
QY 1020 TCCAAACCGGTAGACACAGCACTATCCGCTGCGGCTTATCCGGTAACATATCGCTTTGAG 1079
Db 5422 TCCAAACCGGTAGACACAGCACTTATCCGCTGCGGCTTATCCGGTAACATATCGCTTTGAG 5481
QY 1080 AGACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTAC 1139

Db 5482 AGAGCAGGATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACGCTAC 5541
QY 1140 ACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAA 1199
Db 5542 ACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAA 5601
QY 1200 GTTGGTAGCTCTGTAGTCGGCAAAACAAACACCGTGTGAGCGGTGGTTTTTTTTC 1259
Db 5602 GTTGGTAGCTCTGTAGTCGGCAAAACAAACACCGTGTGAGCGGTGGTTTTTTTTC 5661
QY 1260 AAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
Db 5662 AAGCAGCAGATTACGCGCAGAAAAAAGGATCT 5694

RESULT 8
A60212 A60212 5865 bp DNA circular PAT 06-MAR-1998
LOCUS Sequence 8 from Patent WO9708330.
DEFINITION A60212
ACCESSION A60212
VERSION A60212.1 GI:3715220
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 5865)
AUTHORS Collins, M.K., Weiss, R.A., Takeuchi, Y. and Cosset, F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 8 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES Location/Qualifiers
source 1..5865
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1492 a 1602 c 1390 g 1373 t 8 others
ORIGIN

Query Match 33.4%; Score 604.4; DB 6; Length 5865;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTACTGGCTGTCTCCCTTATCGGGAAGCGGGCGGCATCATATCAATGACGCGCGCTG 654
Db 3752 TTATCATGTCTGGATCCAGATCTGGGCCATCTGGCGCGGATCGATNNNNACATGTGAG 3811
QY 655 TAAAGTTTACGTTGAGAAGAAATTCCTGCAGCCCGCGGTTGCTGGCGTTTTCATATA 714
Db 3812 CAAAAGGCCAGCAAAAGCGGAAACCGTAAAGAGCGCGTTGCTGGCGTTTTCATATA 3871
QY 715 GGTCTCGCCCGCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACC 774
Db 3872 GGTCTCGCCCGCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACC 3931
QY 775 CGACAGGACTATAAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 834
Db 3932 CGACAGGACTATAAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 3991
QY 835 TTCCGACCTTGGCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCG 894
Db 3992 TTCCGACCTTGGCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCG 4051
QY 895 TTCTCAATGCTCACGCTGTAGTATCTCAGTTGCGGTAGGTCGTTCCCTCCCAAGCTGG 954
Db 4052 TTCTCAATGCTCACGCTGTAGTATCTCAGTTGCGGTAGGTCGTTCCCTCCCAAGCTGG 4111
QY 955 GCTGTGTGACGAACCCCGGTTTACGCCGACGCTGCGCTTTATCCGGTAACTATCGTC 1014
Db 4112 GCTGTGTGACGAACCCCGGTTTACGCCGACGCTGCGCTTTATCCGGTAACTATCGTC 4171
QY 1015 TTGAGTCCAAACCGGTTAAGACACGAGCTTATCGCCACTGGCAGCAGCCTGGTAAACAGGA 1074
Db 4172 TTGAGTCCAAACCGGTTAAGACACGAGCTTATCGCCACTGGCAGCAGCCTGGTAAACAGGA 4231

QY 1075 TTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACG 1134
Db 4232 TTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACG 4291
QY 1135 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4292 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4351
QY 1195 AAGAGTTGCTAGCTCTTGTATCCGCAAAACAAACACCGTGTGAGCGGTGGTTTTTTT 1254
Db 4352 AAGAGTTGCTAGCTCTTGTATCCGCAAAACAAACACCGTGTGAGCGGTGGTTTTTTT 4411
QY 1255 TTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
Db 4412 TTTCGAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4449

RESULT 9
AR122288 AR122288 5865 bp DNA PAT 16-MAY-2001
LOCUS Sequence 8 from patent US 6165715.
DEFINITION AR122288
ACCESSION AR122288
VERSION AR122288.1 GI:14106605
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5865)
AUTHORS Collins, M.KatherineLevinge, Weiss, R.Anthony, Takeuchi, Y. and Cosset, F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 8 26-DEC-2000;
FEATURES Location/Qualifiers
source 1..5865
/organism="unknown"
BASE COUNT 1492 a 1602 c 1390 g 1373 t 8 others
ORIGIN

Query Match 33.4%; Score 604.4; DB 6; Length 5865;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTACTGGCTGTCTCCCTTATCGGGAAGCGGGCGGCATCATATCAATGACGCGCGCTG 654
Db 3752 TTATCATGTCTGGATCCAGATCTGGGCCATCTGGCGCGGATCGATNNNNACATGTGAG 3811
QY 655 TAAAGTTTACGTTGAGAAGAAATTCCTGCAGCCCGCGGTTGCTGGCGTTTTCATATA 714
Db 3812 CAAAAGGCCAGCAAAAGCGGAAACCGTAAAGAGCGCGTTGCTGGCGTTTTCATATA 3871
QY 715 GGTCTCGCCCGCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACC 774
Db 3872 GGTCTCGCCCGCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACC 3931
QY 775 CGACAGGACTATAAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 834
Db 3932 CGACAGGACTATAAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 3991
QY 835 TTCCGACCTTGGCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCG 894
Db 3992 TTCCGACCTTGGCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCG 4051
QY 895 TTCTCAATGCTCACGCTGTAGTATCTCAGTTGCGGTAGGTCGTTCCCTCCCAAGCTGG 954
Db 4052 TTCTCAATGCTCACGCTGTAGTATCTCAGTTGCGGTAGGTCGTTCCCTCCCAAGCTGG 4111
QY 955 GCTGTGTGACGAACCCCGGTTTACGCCGACGCTGCGCTTTATCCGGTAACTATCGTC 1014
Db 4112 GCTGTGTGACGAACCCCGGTTTACGCCGACGCTGCGCTTTATCCGGTAACTATCGTC 4171
QY 1015 TTGAGTCCAAACCGGTTAAGACACGAGCTTATCGCCACTGGCAGCAGCCTGGTAAACAGGA 1074

|||||
Db 4172 TTGAGTCCACCCGTAAGACACGACTTATCGCCACTGGCAGCACCCTGGTAACAGGA 4231
QY 1075 TTACAGAGCAGGATGATAGCGGTGTACAGAGTCTTGAAGTGGTGCCCTAACTACG 1134
Db 4232 TTACAGAGCAGGATGATAGCGGTGTACAGAGTCTTGAAGTGGTGCCCTAACTACG 4291
QY 1135 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTACCTTCGGAA 1194
Db 4292 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTACCTTCGGAA 4351
QY 1195 AAAGAGTTGGTACGCTTGTATCGGCAAAACCAACACCGCTGGTAGCGGTGTTTTTTG 1254
Db 4352 AAAGAGTTGGTACGCTTGTATCGGCAAAACCAACACCGCTGGTAGCGGTGTTTTTTG 4411
QY 1255 TTGCAAGCAGCAGATTACGCGCCAGAAAAGGATCT 1292
Db 4412 TTGCAAGCAGCAGATTACGCGCCAGAAAAGGATCT 4449
RESULT 10
A60209 A60209 6028 bp DNA circular PAT 06-MAR-1998
LOCUS
DEFINITION Sequence 5 from Patent WO9708330.
ACCESSION A60209
VERSION A60209.1 GI:3715217
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Collins, M.K., Weiss, R.A., Takeuchi, Y. and Cosset, F.
TITLE EXPRESSION SYSTEMS
JOURNAL PATENT: WO 9708330-A 5 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
source 1. .6028
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1515 a 1637 c 1499 g 1369 t 8 others
ORIGIN
Query Match 33.4%; Score 604.4; DB 6; Length 6028;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGCGCGCTG 654
Db 3915 TTATCATGTCTGGATCCAGATCTGGGCCCATCTGGCGCGGATCGATNNNNACATGTGAG 3974
QY 655 TAAAGTGTACGTTGAGAAGAATTCCTGCAGCCCGCGGTGCTGGCGTTTTCCTATA 714
Db 3975 CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGCGCGTGTGCTGGCGTTTTCCTATA 4034
QY 715 GGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAAC 774
Db 4035 GGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAAC 4094
QY 775 CGAGAGGACTATAAAGATACAGCGGTTCCTCCCTTGGAAAGTCCCTCGCGCTCTCCTG 834
Db 4095 CGAGAGGACTATAAAGATACAGCGGTTCCTCCCTTGGAAAGTCCCTCGCGCTCTCCTG 4154
QY 835 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 894
Db 4155 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 4214
QY 895 TTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAAGCTG 954
Db 4215 TTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAAGCTG 4274
QY 955 GCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
|||||

Db 4275 GCTGTGTGACGAACCCCGCTTCAGCCGACCGCTCGGCTTATCCGGTAACATATCGTC 4334
QY 1015 TTGAGTCCACCCCGTAAAGACACGACTTATCGCCACTGGCAGCACCCTGGTAACAGGA 1074
Db 4335 TTGAGTCCACCCCGTAAAGACACGACTTATCGCCACTGGCAGCACCCTGGTAACAGGA 4394
QY 1075 TTACAGAGCAGGATGATAGCGGTGTACAGAGTCTTGAAGTGGTGCCCTAACTACG 1134
Db 4395 TTACAGAGCAGGATGATAGCGGTGTACAGAGTCTTGAAGTGGTGCCCTAACTACG 4454
QY 1135 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTACCTTCGGAA 1194
Db 4455 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTACCTTCGGAA 4514
QY 1195 AAAGAGTTGGTACGCTTGTATCGGCAAAACCAACACCGCTGGTAGCGGTGTTTTTTG 1254
Db 4515 AAAGAGTTGGTACGCTTGTATCGGCAAAACCAACACCGCTGGTAGCGGTGTTTTTTG 4574
QY 1255 TTGCAAGCAGCAGATTACGCGCCAGAAAAGGATCT 1292
Db 4575 TTGCAAGCAGCAGATTACGCGCCAGAAAAGGATCT 4612
RESULT 11
AR122285
LOCUS AR122285 6028 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6165715.
ACCESSION AR122285
VERSION AR122285.1 GI:14106602
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Collins, M.KatherineLevinge, Weiss, R.Anthony, Takeuchi, Y. and Cosset, F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 5 26-DEC-2000;
FEATURES Location/Qualifiers
source 1. .6028
/organism="unknown"
BASE COUNT 1515 a 1637 c 1499 g 1369 t 8 others
ORIGIN
Query Match 33.4%; Score 604.4; DB 6; Length 6028;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGCGCGCTG 654
Db 3915 TTATCATGTCTGGATCCAGATCTGGGCCCATCTGGCGCGGATCGATNNNNACATGTGAG 3974
QY 655 TAAAGTGTACGTTGAGAAGAATTCCTGCAGCCCGCGGTGCTGGCGTTTTCCTATA 714
Db 3975 CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGCGCGTGTGCTGGCGTTTTCCTATA 4034
QY 715 GGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAAC 774
Db 4035 GGCTCGCCCGCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAAC 4094
QY 775 CGAGAGGACTATAAAGATACAGCGGTTCCTCCCTTGGAAAGTCCCTCGCGCTCTCCTG 834
Db 4095 CGAGAGGACTATAAAGATACAGCGGTTCCTCCCTTGGAAAGTCCCTCGCGCTCTCCTG 4154
QY 835 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 894
Db 4155 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 4214
QY 895 TTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAAGCTG 954
Db 4215 TTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAAGCTG 4274

QY 955 GCTGTGTGCACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACCTATCGTC 1014
Db 4275 GCTGTGTGCACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACCTATCGTC 4334
QY 1015 TTGAGTCCAAACCCGGTAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGA 1074
Db 4335 TTGAGTCCAAACCCGGTAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGA 4394
QY 1075 TTAGCAGAGCAGGTATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACG 1134
Db 4395 TTAGCAGAGCAGGTATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACG 4454
QY 1135 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 1194
Db 4455 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 4514
QY 1195 AAGAGTTGGTAGCTCTGTATCGCGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTG 1254
Db 4515 AAGAGTTGGTAGCTCTGTATCGCGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTG 4574
QY 1255 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db 4575 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 4612
RESULT 12
A60210 A60210 6061 bp DNA circular PAT 06-MAR-1998
LOCUS Sequence 6 from Patent W09708330.
DEFINITION A60210
ACCESSION A60210
VERSION A60210.1 GI:3715218
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 6 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
source
1. .6061
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN
Query Match 33.4%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATAGCGCGCGCTG 654
Db 3948 TTATCATGTCTGGATCCAGATCTGGGCCATCGCGCGCGGATCGATNNNNACATGTGAG 4007
QY 655 TAAAGTGTAGTTGAGAAAGAAATTCGTGAGCGCGCGCGCTTGTGCGGTTCCTCATTA 714
Db 4008 CAAAGGCCAGCAAAAGCCAGAACCCGTAAAGAGCGCGCTTGTGCGGTTCCTCATTA 4067
QY 715 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774
Db 4068 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 4127
QY 775 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCGT 834
Db 4128 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCGT 4187
QY 835 TTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCG 894
Db 4188 TTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCG 4247
QY 895 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTTCCGCTTCCAAAGCTG 954

Db 4248 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTTCCGCTTCCAAAGCTGG 4307
QY 955 GCTGTGTGCACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACCTATCGTC 1014
Db 4308 GCTGTGTGCACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACCTATCGTC 4367
QY 1015 TTGAGTCCAAACCCGGTAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGA 1074
Db 4368 TTGAGTCCAAACCCGGTAAGACACAGCTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGA 4427
QY 1075 TTAGCAGAGCAGGTATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACG 1134
Db 4428 TTAGCAGAGCAGGTATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACG 4487
QY 1135 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 1194
Db 4488 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 4547
QY 1195 AAGAGTTGGTAGCTCTGTATCGCGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTG 1254
Db 4548 AAGAGTTGGTAGCTCTGTATCGCGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTG 4607
QY 1255 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db 4608 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 4645
RESULT 13
AR122286
LOCUS AR122286 6061 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6165715.
ACCESSION AR122286
VERSION AR122286.1 GI:14106603
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.KatherineLevinge, Weiss,R.Anthony, Takeuchi,Y. and Cosset,F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 6 26-DEC-2000;
FEATURES Location/Qualifiers
source
1. .6061
/organism="unknown"
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN
Query Match 33.4%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATAGCGCGCGCTG 654
Db 3948 TTATCATGTCTGGATCCAGATCTGGGCCATCGCGCGCGGATCGATNNNNACATGTGAG 4007
QY 655 TAAAGTGTAGTTGAGAAAGAAATTCGTGAGCGCGCGCGCTTGTGCGGTTCCTCATTA 714
Db 4008 CAAAGGCCAGCAAAAGCCAGAACCCGTAAAGAGCGCGCTTGTGCGGTTCCTCATTA 4067
QY 715 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774
Db 4068 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 4127
QY 775 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCGT 834
Db 4128 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCGT 4187
QY 835 TTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCG 894
Db 4188 TTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCG 4247

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:04:29 ; Search time 599.86 Seconds
(without alignments)
2582.580 Million cell updates/sec

Title: US-09-242-202A-28

Perfect score: 1807
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTCCACGAGCTCTCCACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SID22/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1807	100.0	1807	19	Humanised vector p
2	1583.2	87.6	2308	19	Humanised vector p
3	1310.4	72.5	1547	19	Humanised vector p
c	4	755	41.8	2932	COS cell expressio
c	5	755	41.8	2932	Expression vector
c	6	755	41.8	2932	Expression vector
c	7	755	41.8	2932	Expression vector
c	8	755	41.8	2932	Expression vector
c	9	755	41.8	2932	Eukaryotic cell ex
c	10	755	41.8	2932	Expression vector
11	607.4	33.6	7252	20	Complete sequence

12	604.4	33.4	5865	18	AAT90692	Plasmid FBdelPRDSA	
13	604.4	33.4	6028	18	AAT90689	Plasmid FBdelPASAF	
14	604.4	33.4	6061	18	AAT90690	Plasmid FBdelPMOSA	
15	604.4	33.4	6312	18	AAT90691	Plasmid FBdelPGASA	
16	604.2	33.4	7073	18	AAT50962	TF8-5G9 CDR-grafte	
17	604.2	33.4	7864	18	AAT50963	TF8-5G9 CDR-grafte	
18	604.2	33.4	11795	21	AAA59345	Nucleotide sequenc	
19	604.2	33.4	13254	17	AAT40915	Nucleotide sequenc	
20	604.2	33.4	13254	21	AAA32151	pEel2 Combo BM 12	
c	21	604.2	13254	21	AAA32165	Complete nucleotid	
22	603	33.4	608	19	AAV21731	ColE1 origin of re	
23	603	33.4	2077	20	AAZ23771	Vector pASlib DNA.	
24	603	33.4	2462	21	AAAT74638	Plasmid pSP72. Es	
25	603	33.4	2577	14	AAQ36620	Expression plasmid	
26	603	33.4	2704	20	AAZ06758	Vector pUC28 nucle	
c	27	603	33.4	2927	17	AAT29158	Plasmid pTRP. Syn
c	28	603	33.4	2927	17	AAT31789	Plasmid pTRP. Syn
c	29	603	33.4	2939	16	AAQ87350	Plasmid BGINV. Sy
c	30	603	33.4	2939	19	AAV37293	pBGINV plasmid use
31	603	33.4	3003	11	AAQ05745	Plasmid P_L-mu-smc	
32	603	33.4	3104	10	AAV90296	DNA target sequenc	
33	603	33.4	3130	19	AAV46334	Universal reporter	
c	34	603	33.4	3175	20	AAZ20066	DNA encoding gluta
35	603	33.4	3327	22	AAZ86263	Plasmid GHRH1-29YA	
c	36	603	33.4	3331	20	AAZ84434	pBC66-01, containi
c	37	603	33.4	3331	21	AAZ99245	Nucleotide sequenc
c	38	603	33.4	3344	21	AAA39495	Transgenic unc-119
39	603	33.4	3369	22	AAZ86255	pGHRH-4 44Sk const	
40	603	33.4	3403	9	AAZ80956	Plasmid pDS5/RBSII	
41	603	33.4	3414	12	AAQ12785	pDS56/RBSII-2 sequ	
42	603	33.4	3414	14	AAQ34609	Expression plasmid	
43	603	33.4	3415	12	AAQ12784	pDS56/RBSII-1 sequ	
44	603	33.4	3415	14	AAQ34608	Expression plasmid	
45	603	33.4	3416	9	AAZ80958	Plasmid pDS56/RBSI	

ALIGNMENTS

RESULT 1	
AAV21733	
ID	AAV21733 standard; cDNA; 1807 BP.
XX	
AC	AAV21733;
XX	
DT	17-AUG-1998 (first entry)
XX	
XX	Humanised vector pITL-1.
DE	
XX	Vector; vaccine; tumour; antigen; plasmid pITL-1; ds.
XX	Chimeric - Homo sapiens.
OS	Chimeric - Escherichia coli.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	472..680
FT	/*tag= C
FT	/note= "SupF gene"
FT	686..1292
FT	/*tag= d
FT	/note= "ColE1 origin of replication"
FT	1605..1847
FT	/*tag= e
FT	/note= "RANTES promoter"
XX	
XX	WO9806863-A1.
XX	
PN	Humanised vector p
XX	COS cell expressio
XX	Expression vector
XX	Vector plasmid pIH
PD	Expression vector
XX	Expression vector
XX	Eukaryotic cell ex
PF	Expression vector
XX	97WO-US14306.
XX	
PR	14-AUG-1996; 96US-0023931.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nelson EL, Nelson PJ;
 XX WPI; 1998-159552/14.
 DR Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 PT Claim 14; Page 56; 125pp; English.
 XX Plasmid pVITL-1 comprises a base vector for novel humanised
 CC polynucleotide vectors. The vector in Escherichia coli DH
 CC 10-beta/P3 is deposited as ATCC 98400. Novel vectors comprise a
 CC human-derived promoter or mammalian homologue which is functional
 CC in mammalian target tissue and cells and a sequence acceptance site
 CC (see AAV21735-36) which accepts cDNA products from RT-PCR cloning.
 CC They also contain minimal non-human components, such as a replication
 CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
 CC are necessary for production of the vector. The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX Sequence 1807 BP; 419 A; 463 C; 492 G; 433 T; 0 other;
 SQ

Query Match 100.0%; Score 1807; DB 19; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACTGTCACCATGGCGGATCTTTATCACTGATAGTTGGTGACATATATGTT 60
 DB 1 ggtactgtcacatggcggaattcttatcactgataagttggtggacatatattgtt 60

QY 61 TATCACTGATAAGTGTCAAGCATGACAAAGTTGACGCGAATACAGTATCGTCCGCG 120
 DB 61 tatcactgataaagtgtcaagcatgacaaagtgtgacgcgaatacagtatccgtgcg 120

QY 121 CCTGGACTGTTGAACGAGTGGCGGTAGACGGTCTGACGACACGCAAACTGGCGAAG 180
 DB 121 cctggactgttgaacgaggtgcggtagacggtctgacgacacgcaaaactggcgaa 180

QY 181 GTTGGGGTGACGACGCCGGCGCTTTTACGCACTTCAGGACAAAGCGGCGCTTAAG 240
 DB 181 gttgggggtgacgacgccggcgcttttactggtcacttcaggaaacagggcgccct 240

QY 241 GCATATGTTGAGTGGCTTACCCAGCGGGGATGGGGAGACCTGTAGTCAGAG 300
 DB 241 gccatattggtgagtgagtgcttgcctgacccagcgggagtgaggagacacctgtag 300

QY 301 CCCCAGGCGACGACGCGCAATGCCGCTCTTCCGCTGCAGGATGAGTAGTGCCCTC 360
 DB 301 cccccaggcagcacaggcgaatgccgctcttcccgctccctgcagatgagtagtgcc 360

QY 361 TCTTGGCCCTGGAAGTTGGCACTCAGTGCCGACGAGCTTCTCTTAATAAATTAAG 420
 DB 361 tcttggccctggaagtggcactcagtgccacacagcctgtctcctaataaaattaa 420

QY 421 GCATCATTTTGTCTACAGTGTCTCTATAATATTAAGCTTGTATCGAATTCCTT 480
 DB 421 gcatcattttgtctacagtgtctctataatattaataagcttgcattgcataatct 480

QY 481 CGGACTTTTGAAGTGATGGTGGGGGAGGATTCGAACCTTCGAAGTCGATCGACGC 540
 DB 481 cggacttttgaagtgatgggtgggggaggatttcgaaccttcgaagtcgatacgacgc 540

QY 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCACGGGTAAATGCTTTACT 600
 DB 541 agatttagagtctgctccctttggtccgctcggaacccaccacggttaagtctttact 600

QY 601 GGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATACGCGCGCTGTAAGT 660
 DB 601 ggcctgctcccttatcggaagcgggcgcatcataatcaaatcagcgccgctgtaagt 660

QY 661 GTTACGTTGAGAAAGAAATTCCTGCAGCGCGCTGCTGGCGTTTTCATAGAGCTCC 720
 DB 661 gttacgttgagaaagaattccctgcagccgcgcgtgctg99cgttttcccaagctcc 720

QY 721 GCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAG 780
 DB 721 gccccctgacgagcatcacaaaatcgacgctcaagtgcagagtg99cgaacccgacag 780

QY 781 GACTATAAGATACACAGGCGTTCCCGCTCGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGA 840
 DB 781 gactataagatacaccaggcgtttcccccctggaagctccctcgtgcgtctcctgttccga 840

QY 841 CCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTC 900
 DB 841 cctgctgctttaccggataacctgtccgcgtttctccttcctcg99aagcgtg99cgtttctc 900

QY 901 AATGCTACGCTGTAGGTATCTCAGTTCGCTGAGTTCGCTCCAACTGGGCTGTG 960
 DB 901 aatgctcacgctgtagtattcagttcgggtgtagtgcgttcgctcctcgaagctgggctgtg 960

QY 961 TGCAGCAACCCCGCTACGCGCGCTGCGCTTATCCGCTTATCCGTAACCTCGCTCTTGA 1020
 DB 961 tgcagcaaccccccttcagccgcgcgtgcgtctatcccgtaactcgtctctgagt 1020

QY 1021 CCAACCCGGTAAACACACGACTTATCCGCCACTGCGACGACCACTGTGTAACAGATTAGCA 1080
 DB 1021 ccaacccggttaagacacgacttatcgccactg99cagcagccactg99taacagattagca 1080

QY 1081 GAGCGAGGTATGAGCGGTGTACAGTTCCTGAAGTGGTGGCTTAACAGGCTACA 1140
 DB 1081 gagcgaggtatgtagcggtgtctacagttcttgaagtgggtggcctaactcaggtaca 1140

QY 1141 CTAGAAGGACAGTATTGTTGTTCTGCGCTCTGTAAGCCAGTCTACTCTCGGAAAGAG 1200
 DB 1141 ctagaaggacagtatttggtttctgctgctcgtcgaagccagtctacctc99aaaaagag 1200

QY 1201 TTGGTAGCTCTTATCCGCAACCAACCAACCGCTGCTAGCGGTGTTTGTGTTGCA 1260
 DB 1201 ttggtagctcttgatccgcaacaaacacacccgcgtgtagcgtgtttttttgttgca 1260

QY 1261 AGCAGCAGATTACGCGCAGAAAGATCTGGGGGATCCGGAGAGCTCCCAACGCGTT 1320
 DB 1261 agcagcagattacgcgcagaaaaaaggatctg99ggatcccgagagctcccaacgcgtt 1320

QY 1321 GGATGCATGGATGAGGAAGAGGTAAAGTCTGTAATGAATAAGCAGGAACTTTGAAGA 1380
 DB 1321 ggatgcacggatgaggaagaggttaagatctgtaataagcaggaactttgaaaga 1380

QY 1381 CTCAGTGACTCAGTGAATAAAGACTCAGTACTCTGATCTCTCTCTCTCTCTCTCTCT 1440
 DB 1381 ctctagtgactgaagtaataaagactcagtgactctctgactctctctctctctctct 1440

QY 1441 CTTTGTGTTCCCAAGAAAGCGGCTTCTGCTCTCTGAGGAGGACCCCTTCCCTGGAAGT 1500
 DB 1441 ctttgtgtcccaagaaagcggtctctctgctctctgagggagaccccttccctggaagt 1500

QY 1501 AAAAATAAGGATGTCAGCAGAGAAATTTTCCACCATTGGTGGTGGTAAAGAGAAC 1560
 DB 1501 aaaaataaggatgtcagcagagaaattttccaccattggtgcttggccaagagagaaac 1560

QY 1561 TGATCAGCTCACTCTAGATGAGAGAGCAGTGGAGGAGACAGAGACTCGAATTTCCGGA 1620
 DB 1561 tgatgagctcactctagatgagagagcagtgagggagagacagagactcgaatttccgga 1620


```
QY 1621 GCATTTTCAGTTTCTTTTCCGTTTGTGCAATTTTCACATTATGATACCGGCCCAATGCTTG 1680
Db 1621 gctatttcagtttctttccgttttgcgatttgcgaaatttcacttatgatataccggccaagtcttg 1680
QY 1681 GTTGCTATTGTGAAACTCCCTTAGGGGATGCCCTCAACTGGCCCTATTAAGGGGCCAG 1740
Db 1681 gtgctatttggaaactcccttagggatggccctcaactgagccctataaaggggccag 1740
QY 1741 CTTGAGCTGCAGAGGATTCTGCAGAGGATCAAGACAGCACGTTGGACCTTCGCACAGCCTC 1800
Db 1741 cctgagctgcagaggattctgcagagatcaagacagcacgtggacctgcgacagcctc 1800
QY 1801 TCCACA 1807
Db 1801 tcccaca 1807

RESULT 2
AAV21734
ID AAV21734 standard; cDNA; 2308 BP.
AC AAV21734;
XX
XX 17-AUG-1998 (first entry)
XX Humanised vector pITL-1 GFP.
XX
XX Vector; vaccine; tumour; antigen; plasmid pITL-1 GFP;
KW green fluorescent protein; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Chimeric - Aequorea victoria.
OS Synthetic.
XX
FH Key
FH CDS
FT /*tag= a
FT /product= green fluorescent protein
FT 973..1181
FT /*tag= c
FT /note= "SupF gene"
FT 1191..1793
FT /*tag= d
FT /note= "ColE1 origin of replication"
FT 2063..2308
FT /*tag= e
FT /note= "RANTES promoter"
XX
XX W09806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Example 11; Page 57-58; 125pp; English.
XX
XX Plasmid pITL-1 GFP comprises base vector pITL-1 (see AAV21733) and
XX a humanised green fluorescent protein (GFP) reporter sequence
XX (see AAV21725). Novel humanised vectors comprise a human-derived
XX promoter or mammalian homologue which is functional in mammalian
XX target tissue and cells and a sequence acceptance site which
```

```
CC accepts cDNA products from RT-PCR cloning. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX
SQ Sequence 2308 BP; 571 A; 601 C; 604 G; 532 T; 0 other;

Query Match 87.6%; Score 1583.2; DB 19; Length 2308;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1585; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 220 GAACAAGCGGCGCCTTAAGGGCCATATGTGATGGATCCTTGACCCAGCGGGGAT 279
Db 721 gtacaagtgaagcgcccttaaggccatatgtgagtgccttgacccagcggggat 780
QY 280 GGGGAGACCTGTAGTCAGAGCCCGCGGAGACACAGGCCAATGCCGCTCTTCCCTGC 339
Db 781 gggggagacctgtagtcagagccccggcgagcagccaatgccgctcttccctgc 840
QY 340 AGGATGAGTAGTGAGTGCCTCTCTGGCCTCGGAAGTTCACATCCACAGCCACGCC 399
Db 841 aggatgagtgtgagtgctctctctgcccctggaagtgcctccagtcgccaccagcc 900
QY 400 TTGCTCTAATAAATAAGTTGTCATCATTTTGTCTGACTAGGTGCTCTATATATAT 459
Db 901 ttgtcctaataaataagttgtcatcatatttctgactaggtgtcctcataatat 960
QY 460 AAGCTTGATATCGAATCTTTTCGACTTTTGAAGTGATGTGTGGGGGAAGGATTCGA 519
Db 961 aagcttgatatacgaattcttgcgacttttgaagtgtggtggtgggaagattcga 1020
QY 520 ACCTTCGAAGTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCC 579
Db 1021 accttcgaagtcgatgacggcagatttagagtgctgctcccttggcgctcgggaaaccc 1080
QY 580 ACCACGGGTAAATGCTTTTACTGGCCTGCTCTTATCGGAAGCGGGCGCATCATATCA 639
Db 1081 accacgggtaagtcttttactggcctgtcccttatcggaagcgggcgcatcata 1140
QY 640 AATGACGCGCGCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGCACCCCGCGCTTGC 699
Db 1141 aatgacgcgcgctgtaaagtgttaagttagaagaattcttgacgcccgcgcgttc 1200
QY 700 TGGCGTTTTCATAGGCTCCGCCCTCGACGAGCATCAAAAATCGACGCTCAAGTC 759
Db 1201 tggcgttttccataggctccgcctccctgacgagcatcacaaaaatcgacgctcaagt 1260
QY 760 AGAGGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCTGGAACTCCC 819
Db 1261 agaggtggcgaaacccgacagagactataagataccaggcgtttccctcggaaactcc 1320
QY 820 TCGTGGCTCTCTGTTCCGACCTGCGCGTACCGGATACCTGCGCTTTCCTTCCCTT 879
Db 1321 tcgtgctctctctgttccgacctgcgcttacggatacctgtccgcttctctcctt 1380
QY 880 CGGGAAGCGGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCGTGTAGTCG 939
Db 1381 cgggaaagcggtggcgctttctcaatgctcacgctgtaggtatctcagttcgtgtgagtcg 1440
QY 940 TTCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGCTGCCTTAT 999
Db 1441 ttcgctccaagctgggctgtgtgcagaaaccccccttcagccagcgtgccttat 1500
QY 1000 CCGGTAACTATCGTCTTGTAGTCCAAACCGGTGAGACACAGACTTATCGCTGCGCAG 1059
Db 1501 ccggttaactatgctcttgagttccaaacccggttaagacacgacttatccactcgcagcag 1560
```

[illegible]

CC This COS cell expression vector was constructed from pISV
CC (Little et al., Mol. Biol. Med. 1:473-488 (1983)) by inserting a
CC synthetic transcription unit between the suppressor tRNA gene and
CC the SV40 origin. The transcription unit consisted of a chimeric
CC promoter composed of human cytomegalovirus AD169 immediate early
CC enhancer sequences fused to the HIV LTR -67 to (plus)80 sequences.
CC A polylinker contg. 2 BstXI sites separated by a 350bp stuffer was
CC inserted immediately downstream from the LTR (plus)80 sequence.
CC Downstream from the polylinker were placed the SV40 small
CC antigen splice and early region polyadenylation signals derived
CC from pSV2. The new vector, which forms part of the wider disclosure
CC of the specification, was used in the cloning of the human CD2
CC antigen gene.

XX Sequence 2932 BP; 775 A; 698 C; 582 G; 777 T; 0 other;

Query Match		41.8%;	Score 755;	DB 13;	Length 2932;
Best Local Similarity		97.6%;	Pred. No. 3.3e-210;		
Matches 788;		Conservative 0;	Mismatches 15;	Indels 2;	Gaps 2;
QY	469 ATCGAATCTTCGGACTTTTGAAGTATGATGCTGGTGGGGAAGGATTCGAACCTTCGAA	528			
Db	804 AGCAGATCTTCGCGACTTTTGAAGTATGATGCTGGTGGGGAAGGATTCGAACCTTCGAA	745			
QY	529 GTCGATGACGGCAGATTAGAGTCTCTCCCTTTGGCCGCTCGGAAACCCACACGGGT	588			
Db	744 GTCGATGACGGCAGATTAGAGTCTCTCCCTTTGGCCGCTCGGAAACCCACACGGGT	685			
QY	589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAGCGGGCGCATCATCAATACACGG	648			
Db	584 AATGCTTTTACTGGCTGCTCCCTTATCGGGAGCGGGCGCATCATCAATACACGG	625			
QY	649 CCGCTGTAAGTGTAGCTTGTGAGAAAGATTCTCTGAGCGCCGCTTCTGGCGTTT	708			
Db	624 CCGCTGTAAGTGTAGCTTGTGAGAAAGATTCTCTGAGCGCCGCTTCTGGCGTTT	568			
QY	709 TCCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGGC	768			
Db	567 TCCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCGACCTCAAGTCAGAGTGGC	508			
QY	769 GAAACCCGACAGGACTATAAGATACACAGCGTTTCCCTCGGAAGCTCCCTCGTGGCT	828			
Db	507 GAAACCCGACAGGACTATAAGATACACAGCGTTTCCCTCGGAAGCTCCCTCGTGGCT	448			
QY	829 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCTTCTCCCTCGGGAAGCG	888			
Db	447 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCTTCTCCCTCGGGAAGCG	388			
QY	889 TGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCCA	948			
Db	387 TGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCCA	328			
QY	949 AGCTGGCTGTGTGACGAAACCCCGCTTACGCGCGCGCTTATCCGCTAACT	1008			
Db	327 AGCTGGCTGTGTGACGAAACCCCGCTTACGCGCGCGCTTATCCGCTAACT	268			
QY	1009 ATCGTCTTGTAGTCCACCGGTGTAAGACAGCTTATCGCACTGCGCAGCAGCACTGGTA	1068			
Db	267 ATCGTCTTGTAGTCCACCGGTGTAAGACAGCTTATCGCACTGCGCAGCAGCACTGGTA	208			
QY	1069 ACAGGATTACGAGGAGGATATGATAGCGGTGCTACAGAGTCTTGAAGTGTGGCCTA	1128			
Db	207 ACAGGATTACGAGGAGGATATGATAGCGGTGCTACAGAGTCTTGAAGTGTGGCCTA	148			
QY	1129 ACTAGGCTACACTAGACGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTACT	1188			
Db	147 ACTAGGCTACACTAGACGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTT-CCT	89			
QY	1189 TCGGAAAAAGAGTTGGTAGCTCTTGATTCGGGAAACAAACCCCGCTGTAGCGTGGTT	1248			
Db	88 TCGGAAAAAGAGTTGGTAGCTCTTGATTCGGGAAACAAACCCCGCTGTAGCGTGGTT	29			

QY	1249	TTTTGTGTTGCAAGCAGCAGATTACGC	1275
Db	28	TTTTGTGTTGCAAGCAGCAGATTACGC	2
RESULT	5		
AAT38557/c			
ID	AAT38557	standard; DNA; 2932 BP.	
XX	AAT38557;		
XX	21-JAN-1997	(first entry)	
XX	Expression vector	piH3.	
XX	CD19; cell surface antigen; T lymphocyte; immunoselection cloning;		
KW	vector; piH3; cyclic.		
XX	Synthetic.		
XX	Key	Location/Qualifiers	
FT	misc_feature	1..589	
FT		/tag= a	
FT		/note= "bases derived from pMB1 origin (pBR322	
FT		ori)"	
FT	misc_feature	590..597	
FT		/tag= b	
FT		/function= SacI linker	
FT	misc_feature	598..799	
FT		/tag= c	
FT		/function= synthetic tyrosine suppressor tRNA gene	
FT	LTR	800..947	
FT		/tag= d	
FT		/function= ASV LTR fragment (PvuII to MluI)	
FT	enhancer	948..1500	
FT		/tag= e	
FT		/function= human cytomegalovirus AD169 enhancer	
FT	TATA_signal	1501..1650	
FT		/tag= f	
FT		/function= HIV TATA and tat-responsive elements	
FT	misc_feature	1651..1761	
FT		/tag= g	
FT		/function= pILNAN polylinker (HindIII to Xba)	
FT	misc_signal	1717..2569	
FT		/tag= h	
FT		/function= pSV splice and poly-Addition signals	
FT	rep_origin	2570..2917	
FT		/tag= i	
FT		/function= SV40 origin of replication (PvuII to HindIII)	
FT	misc_feature	2918..2922	
FT		/tag= j	
FT		/function= pIVX, remnant of RI site from polylinker	
XX	EP739980-A2.		
XX	30-OCT-1996.		
XX	23-FEB-1989;	89EP-0103127.	
XX	25-FEB-1988;	88US-0160416.	
XX	(MASS-) MASSACHUSETTS GEN HOSPITAL.		
XX	Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo C;		
PI	Seed B, Simmons D, Stamenkovic I, Stengelin S;		
XX	WPI; 1996-478744/48.		
XX	cDNA encoding CD19 cell surface antigen - useful for recombinant		
XX	antigen prodn. for diagnostic and therapeutic purposes		
PS	Disclosure; Fig 1; 23pp; English.		

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Expression vector piH3 (AAT38557), deposited as ATCC 67634, can be used in a novel immunoselection cloning method for cell surface antigen genes, e.g. human CD19 (see also AAT38556). This involves introducing cDNA fragments into the vector, transfecting mammalian cells in culture, maintaining the cells under conditions that allow antigen expression, exposing the cells to an antibody directed against the antigen, and recovering positive cells by contact with an immobilised antibody directed against the first antibody.

Sequence 2932 BP; 773 A; 700 C; 682 G; 777 T; 0 other;

Query Match 41.8%; Score 755; DB 17; Length 2932;

Best Local Similarity 97.6%; Pred. No. 3.3e-210;

Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTTCGGACTTTTGAAGTATGTGTGGGGAGAGATCGAACCTTCGAA 528

Db 804 AGCAGATTCTTTTCGGACTTTTGAAGTATGTGTGGGGAGAGATCGAACCTTCGAA 745

QY 529 GTCCGATGACGCGAGATTAGAGTCTGCTCCCTTGGCCGCTCGGGAACCCACCGGCT 588

Db 744 GTCCGATGACGCGAGATTAGAGTCTGCTCCCTTGGCCGCTCGGGAACCCACCGGCT 685

QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGGG 648

Db 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGGG 625

QY 649 CGCGTGAAGTGTACGTGTGAGAAGAAATCTCTGAGCCCGCGGTGTCTGGCGTTTT 708

Db 624 CGCGTGAAGTGTACGTGTGAGAAGAAATCTCTGAGCCCGCGGTGTCTGGCGTTTT 568

QY 709 TCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGC 768

Db 567 TCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGC 508

QY 769 GAAACCCGACGAGACTATAAAGATACAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCT 828

Db 507 GAAACCCGACGAGACTATAAAGATACAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCT 448

QY 829 CTCCTGTTCGACCCCTCGCGCTTACCGGATACGTGTCGCGCTTTCTCCCTTCGGGAAGCG 888

Db 447 CTCCTGTTCGACCCCTCGCGCTTACCGGATACGTGTCGCGCTTTCTCCCTTCGGGAAGCG 388

QY 889 TGGCGCTTTCTCAATGCTCACGCTGATGATATCTAGTTCGGTGTAGTTCGCTCCA 948

Db 387 TGGCGCTTTCTCAATGCTCACGCTGATGATATCTAGTTCGGTGTAGTTCGCTCCA 328

QY 949 AGCTGGCTGTGTGCAACACCCCGTTTCAGCCCGACCGCTGCGCTTATCCGGTAACT 1008

Db 327 AGCTGGCTGTGTGCAACACCCCGTTTCAGCCCGACCGCTGCGCTTATCCGGTAACT 268

QY 1009 ATCGTCTTGAGTCCAAACCGGTAAGACACAGCTTATCGCCACTGGCAGCAGCACTGGTA 1068

Db 267 ATCGTCTTGAGTCCAAACCGGTAAGACACAGCTTATCGCCACTGGCAGCAGCACTGGTA 208

QY 1069 ACAGGATTAGCAGAGCGAGTATGTAGGCGGTCTACAGATCTTGAAGTGTGGCCTTA 1128

Db 207 ACAGGATTAGCAGAGCGAGTATGTAGGCGGTCTACAGATCTTGAAGTGTGGCCTTA 148

QY 1129 ACTACGGCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 1188

Db 147 ACTACGGCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT-CCT 89

QY 1189 TCGGAAGAGTGTGGTAGCTCTTGTATCCGGCAACAAACACCGCTGTAGCGGTGGTT 1248

Db 88 TCGGAAGAGTGTGGTAGCTCTTGTATCCGGCAACAAACACCGCTGTAGCGGTGGTT 29

QY 1249 TTTTGTGTTTGAAGCAGCAGATTACGC 1275

Db 28 TTTTGTGTTTGAAGCAGCAGATTACGC 2

RESULT 6		
AAT14702/c		
ID	AAT14702 standard; DNA; 2932 BP.	
XX		
AC	AAT14702;	
XX		
DT	30-OCT-1996 (first entry)	
XX		
DE	Vector plasmid pIH3.	
XX		
KW	Cell surface antigen; cloning; immunoselection; immunotherapy;	
KW	therapy; diagnosis; vector; pIH3; CD2; T-cell receptor; COS;	
KW	T-lymphocyte; ss.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	misc_feature	1..589
FT	/*tag= a	
FT	/note= "bases 1-589 are derived from the pMB1 origin (pBR322 Ori)"	
FT		
FT	misc_feature	590..597
FT	/*tag= b	
FT	/note= "bases 590-597 are derived from the SacII linker"	
FT		
FT	misc_feature	598..799
FT	/*tag= c	
FT	/note= "bases 598-799 are derived from the synthetic tyrosine suppressor tRNA gene (supF gene)"	
FT		
FT	LTR	800..947
FT	/*tag= d	
FT	/note= "bases 800-947 are derived from a remnant of the ASV LTR fragment (PvuII-MluI)"	
FT		
FT	enhancer	948..1500
FT	/*tag= e	
FT	/note= "bases 948-1500 are derived from the hCMV AD169 enhancer"	
FT		
FT	misc_signal	1501..1650
FT	/*tag= f	
FT	/note= "bases 1501-1650 are derived from HIV TATA and tat-responsive elements"	
FT		
FT	misc_feature	1651..1716
FT	/*tag= g	
FT	/note= "bases 1651-1716 are derived from the pLXAN polylinker (HindIII-Xba)"	
FT		
FT	misc_signal	1717..2569
FT	/*tag= h	
FT	/note= "bases 1717-2569 are derived from pSV to splice and poly-Addition signals"	
FT		
FT	misc_feature	2570..2917
FT	/*tag= i	
FT	/note= "bases 2570-2917 are derived from the SV40 origin of replication (PvuII-HindIII)"	
FT		
FT	misc_feature	2918..2922
FT	/*tag= j	
FT	/note= "bases 2918-2922 are derived from pIVX, remnant of R1 site from polylinker"	
FT		
XX	US5506126-A.	
PN		
XX		
PD	09-APR-1996.	
XX		
PF	25-FEB-1988;	88US-0160416.
XX		
PR	01-DEC-1992;	92US-0983647.
PR	25-FEB-1988;	88US-0160416.
PR	13-JUL-1989;	89US-0379076.
PR	13-JUL-1990;	90US-0553759.
PR	18-OCT-1993;	93US-0139273.
XX		
PA	(GEHO) GEN HOSPITAL CORP.	

Fri Jan 18 08:27:46 2002

Db 147 ACTACGGTACACTAGAGCAGTATTGGTATCTGGCTCTGCTGAAGCAGTT-CCT 89
QY 1189 TCGGAAAAAGAGTTGGTAGCTCTTATCCGCAAAACACCGCTGTTAGCGGTGTT 1248
Db 88 TCGGAAAAAGAGTTGGTAGCTCTTATCCGCAAAACACCGCTGTTAGCGGTGTT 29
QY 1249 TTTTGTCTTTCGCAAGCAGCAGATTAGC 1275
Db 28 TTTTGTCTTTCGCAAGCAGCAGATTAGC 2

RESULT 7
AAV63441/C
ID AAV63441 standard; DNA; 2932 BP.
XX AAV63441;
XX 07-JUN-1999 (first entry)
XX Expression vector pIH3.
XX Vector; pIH3; cell surface antigen; lymphocyte; human; cloning;
XX COS; ss.

XX Chimeric - Escherichia coli.
OS Chimeric - Human cytomegalovirus.
OS Chimeric - Human immunodeficiency virus.
OS Chimeric - Rhesus macaque polyoma virus.
XX Key Location/Qualifiers
FT misc_feature 1..589 /*tag= a
FT /*note= "from pMB1 origin (pBR322 ori)"
FT misc_feature 590..597 /*tag= b
FT /*note= "Sali linker"
FT misc_feature 598..799 /*tag= c
FT /*note= "from synthetic tyrosine suppressor tRNA
FT LTR 800..947 /*tag= d
FT /*tag= "ASV ITR fragment"
FT enhancer 948..1500 /*tag= e
FT /*tag= "human cytomegalovirus AD169 enhancer"
FT TATA_signal 1501..1650 /*tag= f
FT /*tag= "HIV TATA and tat-responsive elements"
FT misc_feature 1651..1716 /*tag= g
FT /*tag= "pILNAN polylinker"
FT misc_signal 1717..2569 /*tag= h
FT /*tag= "SV40 small t antigen splice and early
FT /*tag= "region polyA signals from pSV2"
FT misc_feature 2570..2917 /*tag= i
FT /*tag= "SV40 origin of replication"
FT misc_feature 2918..2922 /*tag= j
FT /*tag= "from pIVX, remnant of R1 site from
FT polylinker"

XX US830731-A.
XX 03-NOV-1998.
XX 21-MAY-1997;
XX 01-DEC-1992;
XX 25-FEB-1988;
XX 97US-0861205.
XX 92US-0983647.
XX 88US-0160416.

XX Aruffo A, Seed B;
XX WPI; 1996-200279/20.
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX Example 1; Fig 1A-B; 79pp; English.
XX Vector pIH3 (AAT14702) is a COS cell expression vector constructed
XX from pIVS by inserting a synthetic transcription unit between the
XX suppressor tRNA gene and the SV40 origin. This vector allows
XX the generation of very large mammalian expression libraries and
XX yields large amounts of protein in mammalian host cells, resulting
XX in efficient selection. Efficiency of library construction is
XX comparable to that achieved with phage lambda, but the clones
XX generated are easier to manipulate. A novel immunoselection
XX cloning method was used to clone genes (see also AAT14703-04 and
XX AAT14706-26) for cell surface antigens of human lymphocytes.
XX Sequence 2932 BP; 774 A; 698 C; 683 G; 777 T; 0 other;

Query Match 41.8%; Score 755; DB 17; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.3e-210;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;
QY 469 ATCGATTCTTCGGACTTTTGAAGTGTGTTGGGGAGGATTGCAACCTTCGAA 528
Db 804 AGCAGATCTTCCTCGGACTTTTGAAGTGTGTTGGGGAGGATTGCAACCTTCGAA 745
QY 529 GTGATGACGCGAGATTAGAGTCTGCTCCCTTGGCCCTGGGAAACCCACCGGT 588
Db 744 GTCGATGACGCGAGATTAGAGTCTGCTCCCTTGGCCCTGGGAAACCCACCGGT 685
QY 589 AATGCTTTTACTGGCCCTGCTCCCTTATCGGAAAGCGGCGCATCATCAATGACGG 648
Db 684 AATGCTTTTACTGGCCCTGCTCCCTTATCGGAAAGCGGCGCATCATCAATGACGG 625
QY 649 CCGCTGTAAAGTGTACGTTGAGAAAGATTCTGAGACGCCCGCGGTGCTGGCGTTT 708
Db 624 CCGCTGTAAAGTGTACGTTGAGAAAGATTCTGAGACGCCCGCGGTGCTGGCGTTT 568
QY 709 TCCATAGGCTCCGCCCTGACGAGCATCACAATAATGACGCTCAAGTCAGAGTGGC 768
Db 567 TCCATAGGCTCCGCCCTGACGAGCATCACAATAATGACGCTCAAGTCAGAGTGGC 508
QY 769 GAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCTCGTGGCT 828
Db 507 GAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCTCGTGGCT 448
QY 829 CTCGTTTCCGACCGTGGCGTTTACCGATACCTTCGCGCTTCTCCCTCGGGAAGCG 888
Db 447 CTCGTTTCCGACCGTGGCGTTTACCGATACCTTCGCGCTTCTCCCTCGGGAAGCG 388
QY 889 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGGTTCGCTCCA 948
Db 387 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGGTTCGCTCCA 328
QY 949 AGCTGGGTGTGTGACGAAACCCCGTTTACGCGGACCGCTGCGCTTATCCGTAAC 1008
Db 327 AGCTGGGTGTGTGACGAAACCCCGTTTACGCGGACCGCTGCGCTTATCCGTAAC 268
QY 1009 ATCGCTTTGAGTCCAAACCGGTAAACACGACTTATCCGCACTGGCAGCAGCTGGTA 1068
Db 267 ATCGCTTTGAGTCCAAACCGGTAAACACGACTTATCCGCACTGGCAGCAGCTGGTA 208
QY 1069 ACAGGATTACGAGCGAGGTATGAGCGGTGCTACAGATTCTTGAAGTGGTGGCTTA 1128
Db 207 ACAGGATTACGAGCGAGGTATGAGCGGTGCTACAGATTCTTGAAGTGGTGGCTTA 148
QY 1129 ACTACGGCTACACTAGAAGCAGATTTTGGTATCTGGCTCTGCTGAAGCAGTTACCT 1188

FT /gene= "supf"
FT /note= "synthetic tyrosine suppressor tRNA gene"
FT LTR 800..947
FT /*tag= d
FT /note= "remnant of ASV LTR fragment"
FT enhancer 948..1500
FT /*tag= e
FT /note= "human cytomegalovirus AD169 enhancer"
FT TATA_signal 1501..1650
FT /*tag= f
FT /note= "HIV TATA and tat-responsive element"
FT misc_feature 1651..1716
FT /*tag= g
FT /note= "derived from pILNAN polylinker"
FT polyA_site 1717..2569
FT /*tag= h
FT /note= "derived from pSV to splice and poly-A site"
FT misc_feature 2570..2917
FT /*tag= i
FT /function= "replication origin"
FT /note= "from SV40 origin of replication"
FT misc_feature 2918..2922
FT /*tag= j
FT /note= "from pIVX"
FT US6111093-A.
FT XX
FT PN XX
FT XX
FT PD 29-AUG-2000.
FT XX
FT PF 28-OCT-1998; 98US-0181612.
FT XX
FT PR 01-DEC-1992; 92US-0983647.
FT PR 25-FEB-1988; 88US-0160416.
FT PR 13-JUL-1989; 89US-0379076.
FT PR 23-MAR-1990; 90US-0498809.
FT PR 13-JUL-1990; 90US-0553759.
FT XX
FT PA (GEHO) GEN HOSPITAL CORP.
FT XX
FT PI Stamenkovic I, Seed B;
FT XX
FT DR WPI; 2000-586382/55.
FT XX
FT PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
FT PT useful for immunodiagnosis and immunotherapy of immune-mediated
FT PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
FT PT diseases
FT XX
FT PS Example 1; Fig 1A-B; 75pp; English.
FT CC
FT CC The present sequence is that of COS cell expression vector
FT CC plasmid pIH3. The vector was constructed from pISV by inserting
FT CC a synthetic transcription unit between the suppressor tRNA gene and
FT CC the SV40 origin. The transcription unit consisted of a chimeric
FT CC promoter composed of human cytomegalovirus AD169 immediate early
FT CC enhancer sequences fused to the HIV LTR -67 to +80 sequences.
FT CC Immediately downstream from the LTR +80 was inserted a polylinker
FT CC containing 2 BstXI sites separated by a 350 bp buffer. The BstXI
FT CC sites were flanked by XbaI sites, which could also be used to
FT CC excise the insert. Downstream from the polylinker were placed the
FT CC SV40 small t antigen splice and early region polyA signals derived
FT CC from pSV2. The vector can be used to construct large mammalian
FT CC cDNA expression libraries. A new method for cloning cDNA encoding
FT CC cell surface antigens is based upon transient expression of antigen
FT CC in eukaryotic (e.g. COS) cells and physical selection of cells
FT CC expressing the antigen by adhesion to (panning on) an antibody
FT CC coated substrate such as a culture dish. The purified cDNA and the
FT CC protein that it encodes are useful for immunodiagnostic and
FT CC immunotherapeutic applications, including the diagnosis and
FT CC treatment of immune-mediated infections, diseases, and disorders in
FT CC animals, including humans. These disorders include asthma,
FT CC immune-complex disease, amyloidosis, parasitic diseases or multiple
FT CC sclerosis. High efficiency cDNA expression vectors such as pIH3

CC were designed for use in the novel methods of the invention,
CC yielding large amounts of protein in mammalian host cells,
CC resulting in efficient selection.
XX
SQ Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 41.8%; Score 755; DB 21; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.3e-210;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATTTCTTCGGACTTTTGAAGTGTGTTGGGGGAAGATTTCGAACCTTCGAA 528
DB 804 AGCAGATTCTTTTCGGACTTTTGAAGTGTGTTGGGGGAAGATTTCGAACCTTCGAA 745
QY 529 GTCCATGACGCGACATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACACGCGGT 588
DB 744 GTCATGACGCGACATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACACGCGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGCG 648
DB 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGCG 625
QY 649 CCGCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGACGCCGCCGCTTGTGCGGTTTT 708
DB 624 CCGCTGTAAAGTGTACGTTGAGAAAGA--CCGCGGTAAATTCGCGCTTGTGCGGTTTT 568
QY 709 TCCATAGGCTCCGCCCTCCCTGACGAGCATCAAAATTCGAGCTCAAGTCAGAGTGGC 768
DB 567 TCCATAGGCTCCGCCCTCCCTGACGAGCATCAAAATTCGAGCTCAAGTCAGAGTGGC 508
QY 769 GAAACCGGACAGGACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGCTGCCT 828
DB 507 GAAACCGGACAGGACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGCTGCCT 448
QY 829 CTCCTGTTCCGACCTCGCGCTTACCGGANACCTGTGCGCCTTCTCCCTTCGGGAAGCG 888
DB 447 CTCCTGTTCCGACCTCGCGCTTACCGGANACCTGTGCGCCTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCCA 948
DB 387 TGGCGCTTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCCCA 328
QY 949 AGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGCGAGCTGCGGCTTATCCGGTAAC 1008
DB 327 AGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGCGAGCTGCGGCTTATCCGGTAAC 268
QY 1009 ATCGTCTTTAGTCCAAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACT 1068
DB 267 ATCGTCTTTAGTCCAAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACT 208
QY 1069 ACAGATTAGCAGAGGAGGTATGAGCGGTGTACAGAGTCTTGAAGTGGTGGCCTA 1128
DB 207 ACAGATTAGCAGAGGAGGTATGAGCGGTGTACAGAGTCTTGAAGTGGTGGCCTA 148
QY 1129 ACTAGGCTACACTAGAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCAGTTACT 1188
DB 147 ACTAGGCTACACTAGAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCAGTTACT 89
QY 1189 TCGGAAAAAGATTGGTAGTCTTGTATCCGGCAACAAACACCGCTGGTAGCGGTGTT 1248
DB 88 TCGGAAAAAGATTGGTAGTCTTGTATCCGGCAACAAACACCGCTGGTAGCGGTGTT 29
QY 1249 TTTTGTGTTTGAAGCAGCAGATTACGC 1275
DB 28 TTTTGTGTTTGAAGCAGCAGATTACGC 2

RESULT 10
AAS03171/c
ID AAS03171 standard; DNA; 2932 BP.
XX
AC AAS03171;

(in this case from GALV) and a selectable marker (SM). It is an example of a recombinant expression vector (REV) of the invention, used to create a packaging cell line. The REV of the invention comprise a gene of interest (GOI) and a SM gene. The SM gene is arranged downstream of the GOI and a GOI associated stop codon is spaced from a start codon of the SM gene to ensure that the SM protein is expressed as a result of translation reinitiation. The cell lines are transformed with two REV's, both are replication deficient, one contains the viral gag-pol gene, the other the viral env gene. By using helper constructs, such as the REV's, which are directly selectable and which provide for high expression of the viral gene, high titre retroviral vectors may be obtained. The packaging cell lines are useful for gene therapy. Prior packaging cell lines using full length retroviral genomes as helper genomes were isolated by cotransfecting them with plasmids encoding selectable markers. However, the helper functions can be lost during the passages of the cells in culture and the current packaging systems provide limited titres of infectious retroviral vectors. Co-transfection with a plasmid encoding a SM does not directly select the best gag-pol-env-expressing cells. The new retroviral packaging cell lines overcome these problems.

Query Match	33.4%	Score 604.4;	DB 18;	Length 6312;
Best Local Similarity	91.4%;	Pred. NO. 4.8e-166;		
Matches 638:	Conservative	0;	Mismatches 60;	Indels 0;
	Gaps	0;		

[illegible]

RESULT 15	
AAT90691	AAT90691 standard; DNA; 6312 BP.
XX	
XX	AAT90691;
XX	
XX	05-JAN-1998 (first entry)
XX	Plasmid FBdelPGASAF coding sequence.
DE	
XX	Packaging-deficient construct; viral gag-pol gene; packaging cell line;
XX	moLoney murine leukaemia virus; MoMLV; viral env gene; helper construct;
KW	gene therapy; ss.
XX	
XX	Synthetic.
OS	
XX	WO9708330-A1.
PN	
XX	06-MAR-1997.
PD	
XX	
XX	23-AUG-1996; 96WO-GB02061.
PF	
XX	
XX	23-AUG-1995; 95GB-0017263.
PR	
XX	
XX	(CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA	
XX	
XX	Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
PI	
XX	
XX	WPI; 1997-179287/16.
DR	
XX	Selectable retroviral packaging cell lines and expression constructs
PT	- comprise selectable gene downstream of gene of interest; are
PT	selectable due to the in-efficiency associated with translation
PT	re-initiation
PT	
XX	Claim 13; Fig 11; 79pp; English.
XX	
XX	This sequence represents the recombinant expression plasmid FBdelPASAF.
CC	This sequence is a packaging-deficient construct having a viral env gene
CC	

Fri Jan 18 08:27:46 2002

us-09-242-202a-28.rng

Page 17

Db 4859 ttgcaagcagcagattacgcgcagagaaaaaggatct 4896

Search completed: January 17, 2002, 12:05:08
Job time: 17895 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:55:46 ; Search time 221.34 Seconds
(without alignments)
1848.946 Million cell updates/sec

Title: US-09-242-202A-28
Perfect score: 1807
Sequence: 1 GGTACTGTGCCACCATGGGCGC.....CTCGCACAGCCTCTCCACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	604.4	33.4	5865	4	US-09-011-745-8	Sequence 8, Appli
2	604.4	33.4	6028	4	US-09-011-745-5	Sequence 5, Appli
3	604.4	33.4	6061	4	US-09-011-745-6	Sequence 6, Appli
4	604.4	33.4	6312	4	US-09-011-745-7	Sequence 7, Appli
5	604.2	33.4	13254	1	US-08-276-852-156	Sequence 156, App
6	604.2	33.4	13254	1	US-08-276-853-170	Sequence 170, App
7	604.2	33.4	13254	1	US-08-899-575-156	Sequence 156, App
8	604.2	33.4	13254	1	US-08-899-575-170	Sequence 170, App
9	604.2	33.4	13254	1	US-08-899-575-156	Sequence 156, App
10	604.2	33.4	13254	1	US-08-899-575-170	Sequence 170, App
11	604.2	33.4	13254	5	PCR-US95-08743-156	Sequence 156, App
12	604.2	33.4	13254	5	PCR-US95-08743-170	Sequence 170, App
13	603	33.4	1905	1	US-08-594-469-9	Sequence 9, Appli
14	603	33.4	1905	2	US-08-906-957-9	Sequence 9, Appli
15	603	33.4	2927	2	US-08-941-647A-1	Sequence 1, Appli
16	603	33.4	2939	1	US-08-1151-512-2	Sequence 2, Appli
17	603	33.4	2939	1	US-08-488-015B-2	Sequence 2, Appli
18	603	33.4	2939	3	US-08-814-412-17	Sequence 17, Appl
19	603	33.4	3003	6	5182260-18	Patent No. 5182260
20	603	33.4	3104	1	US-07-415-307A-1	Sequence 1, Appli
21	603	33.4	3104	1	US-08-371-320-1	Sequence 1, Appli
22	603	33.4	3130	4	US-09-038-141-1	Sequence 1, Appli
23	603	33.4	3301	2	US-08-447-430A-42	Sequence 42, Appl
24	603	33.4	3423	2	US-08-447-430A-40	Sequence 40, Appl
25	603	33.4	3474	2	US-08-447-430A-41	Sequence 41, Appl
26	603	33.4	3474	2	US-08-318-837-10	Sequence 10, Appl
27	603	33.4	3585	1	US-08-362-670B-9	Sequence 9, Appli

C 28	603	33.4	3585	3	US-08-333-576C-9	Sequence 9, Appli
C 29	603	33.4	3585	4	US-08-808-324-9	Sequence 9, Appli
C 30	603	33.4	3585	5	PCR-US94-14030A-9	Sequence 9, Appli
C 31	603	33.4	3623	2	US-07-989-847-13	Sequence 13, Appli
C 32	603	33.4	3623	4	US-08-469-411-13	Sequence 13, Appli
C 33	603	33.4	3632	1	US-07-745-382-13	Sequence 13, Appli
C 34	603	33.4	3632	1	US-07-921-848-13	Sequence 13, Appli
C 35	603	33.4	3632	1	US-08-115-680-3	Sequence 3, Appli
C 36	603	33.4	3632	1	US-07-941-372-3	Sequence 3, Appli
C 37	603	33.4	3632	1	US-08-165-301A-13	Sequence 13, Appli
C 38	603	33.4	3632	3	US-08-810-436-13	Sequence 13, Appli
C 39	603	33.4	3632	5	PCR-US93-08247-3	Sequence 3, Appli
C 40	603	33.4	3632	5	PCR-US94-14179-13	Sequence 13, Appli
C 41	603	33.4	3656	1	US-08-232-463-8	Sequence 8, Appli
C 42	603	33.4	3688	1	US-08-232-463-9	Sequence 9, Appli
C 43	603	33.4	3789	1	US-09-075-019-8	Sequence 8, Appli
C 44	603	33.4	3803	1	US-07-623-953-1	Sequence 1, Appli
C 45	603	33.4	3822	3	US-08-675-566-8	Sequence 8, Appli

ALIGNMENTS

```

RESULT 1
US-09-011-745-8
Sequence 8, Application US/09011745.
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: Collins, Mary KL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Lolc
TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02062
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 5865
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial
OTHER INFORMATION: construct
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3611)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3612)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3613)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3614)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: (3799)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3800)
OTHER INFORMATION: n is any nucleotide

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (3801)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3802)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-8

Query Match 33.4%; Score 604.4; DB 4; Length 5865;
Best Local Similarity 91.4%; Pred. No. 8.8e-175;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGCCCTGCTCCCTTATFCGGAGAGCGGGGCGCATCATCAAAATGACGCCCGCTG 654
DB 3752 ttatcatgtctggatccagatctggcccatcgccgcgagatgcatnnnnacatgtgag 3811
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGACGCCCGCGGTGCTGGCGTTTTCCTG 714
DB 3812 caaagccagcaaaagccaggaacccgtaaaagccgctgctgctgctgtttccata 3871
QY 715 GGCCTCCGCCCTCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAAC 774
DB 3872 ggcctccgccccctgacagatcacaaaaatcgacgctcaagtcagagtggtgcaaac 3931
QY 775 GCACAGACTATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTGCTGCTCCTG 834
DB 3932 cgacagactataaagataacagcgcttccctcggaagctccctgctgctctctg 3991
QY 835 TTCGACCCCTCCGCTTACCGGATACCTGTCGCCCTTTCCTCCCTGCGGAAGCGTGGCG 894
DB 3992 ttcgacccctccgcttaccggaatactgctccgcttctcccttcggaaggtggcg 4051
QY 895 TTTCTCAATGCTCACGCTGTAGTATCTAGTTCGGTGTAGTGTCTCGTCCCAAGCTG 954
DB 4052 ttctcaatgctcaagctgtaggtatctcaagttcggtgagttcgcttcccaagctgg 4111
QY 955 GCTGTGTGACGAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
DB 4112 gctgtgtgacgaaccccgcttaccggaatactgctccgcttctcccttcggaaggtggcg 4051
QY 1015 TTGAGTCCAAACCGGTAAGACAGAGCTTATCGCCACTGCGCAGCAGCAGCTGTAAACAG 1074
DB 4172 ttgagtcgaaccccggttaagacagagcttaccggaatactgctccgcttctcccttcggaaggtggcg 4051
QY 1075 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAGTGGTGGCTTAACAG 1134
DB 4232 ttgacagagcgaggtatgtagcggtgctacagagttcttgaagtggtggtggttaactacg 4291
QY 1135 GCTACACTAGAGGACAGTATTTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
DB 4292 gctacactagaagagcagatttggatctgctgctgctgctgctgctgctgctgctgctg 4351
QY 1195 AAAGAGTTGGTGTGCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGTGTTTTTTG 1254
DB 4352 aaagagttggtgctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 4411
QY 1255 TTGCAAGCAGCAGATATAGCGCAGCAAAAGATGATCT 1292
DB 4412 ttgcaagcagcagattatcgccagcaaaaagagatct 4449

RESULT 2

US-09-011-745-5
Sequence 5, Application US/09011745
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: Collins, Mary KL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Loic

TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 6028
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Portion of
OTHER INFORMATION: construct
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3774)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3775)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3776)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3777)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3962)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3963)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3964)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3965)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-5

Query Match 33.4%; Score 604.4; DB 4; Length 6028;
Best Local Similarity 91.4%; Pred. No. 8.9e-175;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGCCCTGCTCCCTTATFCGGAGAGCGGGGCGCATCATCAAAATGACGCCCGCTG 654
DB 3915 ttatcatgtctggatccagatctggcccatcgccgcgagatgcatnnnnacatgtgag 3974
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGACGCCCGCGGTGCTGGCGTTTTCCTG 714
DB 3975 caaagccagcaaaagccaggaacccgtaaaagccgctgctgctgctgctgctgctgctg 4034
QY 715 GGCCTCCGCCCTCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAAC 774
DB 4035 ggcctccgccccctgacagatcacaaaaatcgacgctcaagtcagagtggtgcaaac 4094
QY 775 GCACAGACTATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTGCTGCTGCTGCTG 834
DB 4095 cgacagactataaagataccagcgcttccctcggaagctccctgctgctgctgctgctg 4154
QY 835 TTCGACCCCTCCGCTTACCGGATACCTGTCCGCTTTCCTGCGGAAGCGTGGCGC 894

```

|||||
Db 4155 ttccgaccctccgcttacggataacctgtccgcctttcccttccttcggaagcgtggcgc 4214
Qy 895 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGCTGTAGGTGTTCGCTCCAGCTGG 954
Db 4215 ttctcaatgctcacgctgtaggatctcagttcggtaggttcgctccaagctgg 4274
Qy 955 GCTGTGTGACGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACATATCGTC 1014
Db 4275 gctgtgtgcagcaaccccgcttcagcccgaccgctgcgcttaccggttaactatcgtc 4334
Qy 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATGCCACTGTGGCAGCCACCTGGTAACAGGA 1074
Db 4335 ttgagtcacaacccggtaagacacgacttatcgccactggcagcagcactggttaacagga 4394
Qy 1075 TTACGAGCGCAGGTATCTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCTAACTACG 1134
Db 4395 ttgacagcgaggtatgtaggcgtgctacagagttcttgaagtggtggcctaactacg 4454
Qy 1135 GCTACACTAGAAGGACAGTATTTGGTATCTGCGTCTGCTGAAGCCAGTTTACCTTCGGAA 1194
Db 4455 gctacactagaagacagatttggtagtctgctgtcgtgtaagccaggttaccttcggaa 4514
Qy 1195 AAAGAGTGTGAGTCTTGTATCCGGCAAAACAAACACCGCTGTGAGCGGTGGTTTTTTG 1254
Db 4515 aaagagttggtagctcttgatcccggaacaaacaccgcgtgtagcggtggttttttg 4574
Qy 1255 TTGCAACGACAGATACGCGCGAGAAAAAAGGATCT 1292
Db 4575 ttgcaagcagcagattacgcgcgcaaaaaaaggatct 4612

```

RESULT 3

```

US-09-011-745-6
; Sequence 6, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 6061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; FEATURE:
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3807)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3808)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3809)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3810)

```

```

; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3995)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3996)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3997)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3998)
; OTHER INFORMATION: n is any nucleotide
; US-09-011-745-6

Query Match      33.4%  Score 604.4;  DB 4;  Length 6061;
Best Local Similarity 91.4%;  Pred. No. 8.9e-175;
Matches 636;  Conservative 0;  Mismatches 60;  Indels 0;  Gaps 0;

Qy 595 TTTACTGCGCTGCTCCCTTATCGGGAGCGGGGCGCATATATCAAAATGACGCGCGCTG 654
Db 3948 ttatcagtctggatccagatctggcccatcgccgcgagatcgatnnnnacatgtgag 4007
Qy 655 TAAAGTGTGTAGTTGAGAAATTCCTGTGACGCCGCGCTTGTGCGGTGTTTTCCATA 714
Db 4008 caaaagccagcaaaagccaggaaacgcgtaaaaagccgcgttgctgctgttttccata 4067
Qy 715 GGCTCCGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774
Db 4068 ggcctccgccccctgacgagcatcaaaaaatcgacgtcaagtcagaggtggcgaaacc 4127
Qy 775 CGACAGGACTATAAAGATACACAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCGT 834
Db 4128 cgacagactataaagataaccagcggtttcccccctggaagctccctcgctcgctctcg 4187
Qy 835 TTCGACCGCTCCCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGGAAGCGTGGCGC 894
Db 4188 ttccgacctgcgcttaccggatcacctgtccgccttctcccttcgggaagcgtggcgc 4247
Qy 895 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTGCTTCGCTCCAGCTGG 954
Db 4248 ttctcaatgctcacgctgtaggatctcagttcgggtgtaggtcgttcgctccaagctgg 4307
Qy 955 GCTGTGTGACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCCGTAACTATCGTC 1014
Db 4308 gctgtgtgcagcaaccccgcttcagcccgaccttcagcccgcttcagcgttaactatcgtc 4367
Qy 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 1074
Db 4368 ttgagtcacaacccggtaagacacgacttatcgccactggcagcagcactggttaacagga 4427
Qy 1075 TTACGAGCGCAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCTAACTACG 1134
Db 4428 ttgacagagcgaggtatgtcagggcgtgctacagagttcttgaagtggtggcctaactacg 4487
Qy 1135 GCTACACTAGAAGGACAGTATTTGGTATCTGCGTCTGCTGAAGCCAGTTTACCTTCGGAA 1194
Db 4488 gctacactagaagacagatttggtagtctgctgctgtaagccaggtacccttcggaa 4547
Qy 1195 AAAGAGTGTGTAGTCTTGTATCCCGGCAAAACAAACACCGCTGTGAGCGGTGTTTTTTG 1254
Db 4548 aaagagttggtagctcttgatcccgcaaaacaaacccgcgtgtagcggtggtgtttttg 4607
Qy 1255 TTGCAACGACAGATACGCGCGAGAAAAAAGGATCT 1292
Db 4608 ttgcaagcagcagattacgcgcgcaaaaaaaggatct 4645

```



```

; PRIOR APPLICATION DATA.:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-276-852-156

Query Match 33.4%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGATTCTGACGCCGCCGCTTCTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC 727
Db 5976 TAAGGGAGAGCGTCGACCTTCGGGGCGCGTCTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC 6035

QY 728 TGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATA 787
Db 6036 TGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATA 6095

QY 788 AAGATACCAGCGTTTCCCCCTGGAAGTCCCTCGTGGCGCTCTCCTGTTCCGACCCCTGCC 847
Db 6096 AAGATACCAGCGTTTCCCCCTGGAAGTCCCTCGTGGCGCTCTCCTGTTCCGACCCCTGCC 6155

QY 848 GCTTACCGGATACCTGTCGGCGTTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTC 907
Db 6156 GCTTACCGGATACCTGTCGGCGTTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTC 6215

QY 908 ACGTGTAGGTATCTCAGTTCGGTGTAGTCTGTCGCTCCAGCTGGGCTGTGTGCACGA 967
Db 6216 ACGTGTAGGTATCTCAGTTCGGTGTAGTCTGTCGCTCCAGCTGGGCTGTGTGCACGA 6275

QY 968 ACCCCCGCTTCAGCCCGCACCCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAAGCC 1027
Db 6276 ACCCCCGCTTCAGCCCGCACCCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAAGCC 6335

QY 1028 GGTAAAGACACACTTATCGCACCTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 1087
Db 6336 GGTAAAGACACACTTATCGCACCTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 6395

QY 1088 GTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCGCTAACTACGCGTACACTAGAAG 1147
Db 6396 GTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCGCTAACTACGCGTACACTAGAAG 6455

QY 1148 GACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAG 1207
Db 6456 GACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAG 6515

QY 1208 CTCATTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTGTGCAAGCAGCA 1267
Db 6516 CTCATTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTGTGCAAGCAGCA 6575

QY 1268 GATTACGGCGCAGAAAAAAGGATCT 1292
Db 6576 GATTACGGCGCAGAAAAAAGGATCT 6600

RESULT 6
US-08-276-852-170/c
; Sequence 170, Application US/08276852
; Patent No. 5652138

```

```

> GENERAL INFORMATION:
> APPLICANT: Burton, Dennis R
> APPLICANT: Barbas, Carlos F
> APPLICANT: Lerner, Richard A
> TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
> TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
> NUMBER OF SEQUENCES: 170
> CORRESPONDENCE ADDRESS:
> ADDRESSEE: The Scripps Research Institute, Office of
> ADDRESSEE: Patent Counsel
> STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
> STREET: Mail Drop TPC8
> CITY: La Jolla
> STATE: CA
> COUNTRY: USA
> ZIP: 92037
> COMPUTER READABLE FORM:
> MEDIUM TYPE: Floppy disk
> OPERATING SYSTEM: PC-DOS/MS-DOS
> SOFTWARE: PatentIn Release #1.0, Version #1.25
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER: US/08/276,852
> FILING DATE: 18-JUL-1994
> CLASSIFICATION: 514
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: US 08/178,302
> FILING DATE: 30-SEP-1993
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: US 07/954,148
> FILING DATE: 30-SEP-1992
> ATTORNEY/AGENT INFORMATION:
> NAME: Fitting, Thomas
> REGISTRATION NUMBER: 34,163
> REFERENCE/DOCKET NUMBER: SCR1452P
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 619-554-2937
> TELEFAX: 619-554-6312
> INFORMATION FOR SEQ ID NO: 170:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 13254 base pairs
> TYPE: nucleic acid
> STRANDEDNESS: double
> TOPOLOGY: circular
> MOLECULE TYPE: DNA (genomic)
> US-08-276-852-170

Query Match          33.4%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps

Qy 668 TGAGAAAGAATTCTCGACGCCGCGGTGTGGCGTTTTCATAGGCTCCGCCCCC 727
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7279 TAAGGGAGAGCGTCGACCTCGGCCGCGTGTGGCGTTTTTTCATAGGCTCCGCCCCC 7220

Qy 728 TGACGACCATCACAAAATCGAGCCTCAAGTCAGAGGTGGCGAAACCCGCACAGGACTATA 787
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7219 TGACGACCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGCACAGGACTATA 7160

Qy 788 AAGATACCAGGCGTTTCCCCCTGGGAAGCTCCCTCGTGCGCTCTCCTGTTCGACCCCTGCC 847
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7159 AAGATACCAGGCGTTTCCCCCTGGGAAGCTCCCTCGTGCGCTCTCCTGTTCGACCCCTGCC 7100

Qy 848 GCTTACCGGATACGTTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGTCTC 907
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7099 GCTTACCGGATACGTTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGTCTC 7040

Qy 908 ACGETGTAGTATCTCAGTTTCGGGTAGTCTCGTCTCAAGCTGGGCTGTGTCCACGA 967
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7039 ACGETGTAGTATCTCAGTTTCGGGTAGTCTCGTCTCAAGCTGGGCTGTGTCCACGA 6980

Qy 968 ACCCCCGGTTTCAGCCCGACCGCTGCGCCCTTATCCGGGTAACATATCGTCTTAGTCCAACCC 1027

```

RESULT . 6
US-08-276-852-170/c
; Sequence 170, Application US/08276852
; Patent No. 5652138

QY 728 TCACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATA 787
Db 6036 TCACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATA 6095
QY 788 AAGATACACAGCGGTTCCCGCTTGGAGCTCCCTCGTGGGCTCTCTGTTCGACCTGCC 847
Db 6096 AAGATACACAGCGGTTCCCGCTTGGAGCTCCCTCGTGGGCTCTCTGTTCGACCTGCC 6155
QY 848 GCTTACCGGATACCTGTGCGGCTTCTCCCTTCGGAAGCGTGGGCTTCTCAATGCTC 907
Db 6156 GCTTACCGGATACCTGTGCGGCTTCTCCCTTCGGAAGCGTGGGCTTCTCAATGCTC 6215
QY 908 ACGCTGTAGTATCTCAGTTCGCTGTAGTGTCTCGCTCAAGCTGGGCTGTGTCACGA 967
Db 6216 ACGCTGTAGTATCTCAGTTCGCTGTAGTGTCTCGCTCAAGCTGGGCTGTGTCACGA 6275
QY 968 ACCCGCGCTTACGCGGCTTCCCGCTTAACTATCGGTAACTATCGTTCGATCCAAACC 1027
Db 6276 ACCCGCGCTTACGCGGCTTCCCGCTTAACTATCGGTAACTATCGTTCGATCCAAACC 6335
QY 1028 GGTAAACAGAGCTTATCCGCTTGGCAGCAGCCTGCTAAACAGGATTAGCAGAGCGAG 1087
Db 6336 GGTAAACAGAGCTTATCCGCTTGGCAGCAGCCTGCTAAACAGGATTAGCAGAGCGAG 6395
QY 1088 GTATGTAGCGGCTGTACAGAGTCTTGAAGTGGTGGCCTAACTAGGCTACACTAGAAG 1147
Db 6396 GTATGTAGCGGCTGTACAGAGTCTTGAAGTGGTGGCCTAACTAGGCTACACTAGAAG 6455
QY 1148 GACAGTATTTGGTATCTGCTGCTGCTGAAGCAGTACTTCCGGAAGAGTTGGTAG 1207
Db 6456 GACAGTATTTGGTATCTGCTGCTGCTGAAGCAGTACTTCCGGAAGAGTTGGTAG 6515
QY 1208 CTCTTCATCCGCAAAACCAACCCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCA 1267
Db 6516 CTCTTCATCCGCAAAACCAACCCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCA 6575
QY 1268 GATTACGCGCAGAAAAAAGGATCT 1292
Db 6576 GATTACGCGCAGAAAAAAGGATCT 6600

RESULT 10
US-08-899-575-170/c
; Sequence 170, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994

APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 33.4%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAATTCCTGCGACCGCGCGCTGTGCTGGCGTTTTCATAGGCTCCGCCGCC 727
Db 7279 TAAGGAGAGCGCTCGACCTCGGGCGCGTGTGCTGGCGTTTTCATAGGCTCCGCCGCC 7220
QY 728 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATA 787
Db 7219 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGGACTATA 7160
QY 788 AAGATACACAGCGGTTTCCCGCTTGGAGCTCCCTCGTGGCTCTCTCTGTTCGACCTGCC 847
Db 7159 AAGATACACAGCGGTTTCCCGCTTGGAGCTCCCTCGTGGCTCTCTCTGTTCGACCTGCC 7100
QY 848 GCTTACCGGATACCTGTGCGGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTC 907
Db 7099 GCTTACCGGATACCTGTGCGGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTC 7040
QY 908 ACGCTGTAGTATCTCAGTTCGCTGTAGTGTCTCGCTCAAGCTGGGCTGTGTCACGA 967
Db 7039 ACGCTGTAGTATCTCAGTTCGCTGTAGTGTCTCGCTCAAGCTGGGCTGTGTCACGA 6980
QY 968 ACCCGCGCTTACGCGGCTTCCCGCTTAACTATCGGTAACTATCGTTCGATCCAAACC 1027
Db 6979 ACCCGCGCTTACGCGGCTTCCCGCTTAACTATCGGTAACTATCGTTCGATCCAAACC 6920
QY 1028 GGTAAACAGAGCTTATCCGCTTGGCAGCAGCCTGCTAAACAGGATTAGCAGAGCGAG 1087
Db 6919 GGTAAACAGAGCTTATCCGCTTGGCAGCAGCCTGCTAAACAGGATTAGCAGAGCGAG 6860
QY 1088 GTATGTAGCGGCTGTACAGAGTCTTGAAGTGGTGGCCTAACTAGCGCTACACTAGAAG 1147
Db 6859 GTATGTAGCGGCTGTACAGAGTCTTGAAGTGGTGGCCTAACTAGCGCTACACTAGAAG 6800
QY 1148 GACAGTATTTGGTATCTGCTGCTGCTGAAGCAGTTCCTTCGGAAGAGTTGGTAG 1207
Db 6799 GACAGTATTTGGTATCTGCTGCTGCTGAAGCAGTTCCTTCGGAAGAGTTGGTAG 6740
QY 1208 CTCTTCATCCGCAAAACCAACCCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCA 1267
Db 6739 CTCTTCATCCGCAAAACCAACCCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCA 6680
QY 1268 GATTACGCGCAGAAAAAAGGATCT 1292
Db 6679 GATTACGCGCAGAAAAAAGGATCT 6655

RESULT 11
PCT-US95-08743-156

Db 247 TTTCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCG 306
QY 930 GTGTAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCCGACCGC 989
Db 307 GTGTAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCCGACCGC 366
QY 990 TGGCGCTTATCCGTAAGTATCTGTTGAGTCCAAACCCCGGTAGACACACTTATCGCCA 1049
Db 367 TGGCGCTTATCCGTAAGTATCTGTTGAGTCCAAACCCCGGTAGACACACTTATCGCCA 426
QY 1050 CTGCGACGACGCTGTTGTAACAGATTAGCAGCGGAGGTATGTAGCGGTGTACAGAG 1109
Db 427 CTGCGACGACGCTGTTGTAACAGATTAGCAGCGGAGGTATGTAGCGGTGTACAGAG 486
QY 1110 TTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGGCT 1169
Db 487 TTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGGCT 546
QY 1170 CTGCTGAAGCCAGTTACTTTCGGAAGAAAGAGTTGGTAGCTCTTGTATCCGCAACAAACC 1229
Db 547 CTGCTGAAGCCAGTTACTTTCGGAAGAAAGAGTTGGTAGCTCTTGTATCCGCAACAAACC 606
QY 1230 ACCGCTGTAGCGGTGTTTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGA 1289
Db 607 ACCGCTGTAGCGGTGTTTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGA 666
QY 1290 TCT 1292
Db 667 TCT 669
RESULT 14
US-08-906-957-9
; Sequence 9, Application US/08906957
; Patent No. 5856142
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; TITLE OF INVENTION: presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,957
; FILING DATE: 06-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/594,469
; FILING DATE:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:

Db 6739 CTTCTGATCCGCAACAAACACCGCTGGTAGCGGTGTTTGTGTTGCAAGCAGCA 6680
QY 1268 GATTAGCGCGCAGAAAAAGGATCT 1292
Db 6679 GATTAGCGCGCAGAAAAAGGATCT 6655
RESULT 13
US-08-594-469-9
; Sequence 9, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; TITLE OF INVENTION: presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,469
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-594-469-9
Query Match 33.4%; Score 603; DB 1; Length 1905;
Best Local Similarity 100.0%; Pred. No. 1.3e-174;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCGCGCTGCTGGCGTTTTCATAGGTCGCCGCCCTGAGCAGCATCACAAAATCGA 749
Db 67 GCGCGCTGCTGGCGTTTTCATAGGTCGCCGCCCTGAGCAGCATCACAAAATCGA 126
QY 750 CGTCAAGTCAGAGTGGCGAACCAGCAGACTATAAGATACAGCGTTTCCCGCT 809
Db 127 CGTCAAGTCAGAGTGGCGAACCAGCAGACTATAAGATACAGCGTTTCCCGCT 186
QY 810 GGAAGCTCCCTCGTGGCGTCTCTGTTCCGACCGCTCGCGCTTACCGGATACCTGTCGCCG 869
Db 187 GGAAGCTCCCTCGTGGCGTCTCTGTTCCGACCGCTCGCGCTTACCGGATACCTGTCGCCG 246
QY 870 TTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCG 929

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-906-957-9

Query Match 33.4%; Score 603; DB 2; Length 1905;
Best Local Similarity 100.0%; Pred. No. 1.3e-174;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCCGGTGTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 749
DB 67 GCCCGGTGTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 126

QY 750 CGCTCAAGTCAGAGGTGGCAACCCGACAGGACTATAAAGATACCAAGGGTTTCCCGCT 809
DB 127 CGCTCAAGTCAGAGGTGGCAACCCGACAGGACTATAAAGATACCAAGGGTTTCCCGCT 186

QY 810 GGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCGC 869
DB 187 GGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCGC 246

QY 870 TTTCTCCCTCGGGAAGGTGGCGTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGC 929
DB 247 TTTCTCCCTCGGGAAGGTGGCGTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGC 306

QY 930 GTGAGGTGCTGCTCCAAAGTGGGCTGTGTGCACGAACCCCGCTTACGCGCGACCGC 989
DB 307 GTGAGGTGCTGCTCCAAAGTGGGCTGTGTGCACGAACCCCGCTTACGCGCGACCGC 366

QY 990 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAACAGGACTTATCGCCA 1049
DB 367 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAACAGGACTTATCGCCA 426

QY 1050 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGAGGTATGAGCGGTGTACAGAG 1109
DB 427 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGAGGTATGAGCGGTGTACAGAG 486

QY 1110 TTTCTGAAGTGGTGGCTTAAGTACGGCTTACAGAGGAGGATTTGGTATCTGCGCT 1169
DB 487 TTTCTGAAGTGGTGGCTTAAGTACGGCTTACAGAGGAGGATTTGGTATCTGCGCT 546

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTTGTATCGGCAACCAACC 1229
DB 547 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTTGTATCGGCAACCAACC 606

QY 1230 ACCGCTGTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACGCGCAGAAAAAGGA 1289
DB 607 ACCGCTGTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACGCGCAGAAAAAGGA 666

QY 1290 TCT 1292
DB 667 TCT 669

RESULT. 15

US-08-941-647A-1/c
; Sequence 1, Application us/08941647A
; Patent No. 5952211
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, ATSUGO
; APPLICANT: TANAKA, TOSHIO
; APPLICANT: MATSUO, YUSHI
; APPLICANT: TANASE, SUMIO
; APPLICANT: FUNATSU, MASAHICO
; APPLICANT: ETO, AKIRA
; TITLE OF INVENTION: METHOD FOR PRODUCING ACTIVE HUMAN
; TITLE OF INVENTION: ALANINE AMINOTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,647A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/533,717
; FILING DATE: 26-SEP-1995
; APPLICATION NUMBER: JP 6-268119
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7335-001-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2927 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-941-647A-1

Query Match 33.4%; Score 603; DB 2; Length 2927;
Best Local Similarity 100.0%; Pred. No. 1.6e-174;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCCGGTGTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 749
DB 2763 GCCCGGTGTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 2704

QY 750 CGCTCAAGTCAGAGGTGGCAACCCGACAGGACTATAAAGATACCAAGGGTTTCCCGCT 809
DB 2703 CGCTCAAGTCAGAGGTGGCAACCCGACAGGACTATAAAGATACCAAGGGTTTCCCGCT 2644

QY 810 GGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCGC 869
DB 2643 GGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCGC 2584

QY 870 TTTCTCCCTCGGGAAGGTGGCGTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGC 929
DB 2583 TTTCTCCCTCGGGAAGGTGGCGTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGC 2524

QY 930 GTGAGGTGCTTCCCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCGCGACCGC 989
DB 2523 GTGAGGTGCTTCCCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCGCGACCGC 2464

QY 990 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAACAGGACTTATCGCCA 1049
DB 2463 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAACAGGACTTATCGCCA 2404

QY 1050 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGAGGTATGAGCGGTGTACAGAG 1109
DB 2403 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGAGGTATGAGCGGTGTACAGAG 2344

QY 1110 TTTCTGAAGTGGTGGCTTAAGTACGGCTTACAGAGGAGGATTTGGTATCTGCGCT 1169
DB 2343 TTTCTGAAGTGGTGGCTTAAGTACGGCTTACAGAGGAGGATTTGGTATCTGCGCT 2284

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACC 1229
Db 2283 CTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGCGCAACAAACC 2224
QY 1230 ACCGCTGGTAGCGGTGGTTTTTTTGTTCAGCAGCAGATTACGCGCAGAAAAAGGA 1289
Db 2223 ACCGCTGGTAGCGGTGGTTTTTTTGTTCAGCAGCAGATTACGCGCAGAAAAAGGA 2164
QY 1290 TCT 1292
Db 2163 TCT 2161

Search completed: January 17, 2002, 11:57:05
Job time: 17552 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 15:14:01 ; Search time 11375.6 Seconds
(without alignments)
2784.055 Million cell updates/sec

Title: US-09-242-202A-28
Perfect score: 1807
Sequence: 1 GGTACCTGCCACCTGGCGC.....CTCGCACAGCCTCTCCACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 17159718 seqs, 8763200856 residues
Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1:	/cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2:	/cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3:	/cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4:	/cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5:	/cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6:	/cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7:	/cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8:	/cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9:	/cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10:	/cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11:	/cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12:	/cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13:	/cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14:	/cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15:	/cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16:	/cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17:	/cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18:	/cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19:	/cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
20:	/cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
21:	/cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22:	/cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23:	/cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
24:	/cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25:	/cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26:	/cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27:	/cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28:	/cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29:	/cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30:	/cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31:	/cgn2_6/ptodata/2/pna/US098_COMB.seq:*
32:	/cgn2_6/ptodata/2/pna/US099_COMB.seq:*
33:	/cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
34:	/cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
35:	/cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
36:	/cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
37:	/cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
38:	/cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
39:	/cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
40:	/cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
41:	/cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
42:	/cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
43:	/cgn2_6/ptodata/2/pna/US6010_COMB.seq:*

44:	/cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
45:	/cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
46:	/cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
47:	/cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
48:	/cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
49:	/cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
50:	/cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
51:	/cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
52:	/cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
53:	/cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
54:	/cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
55:	/cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
56:	/cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
57:	/cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
58:	/cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
59:	/cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
60:	/cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
61:	/cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
62:	/cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
63:	/cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
64:	/cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
65:	/cgn2_6/ptodata/2/pna/US6032_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1807	100.0	1807	16	US-09-242-202-28	Sequence 28, Appl
2	1807	100.0	1807	16	US-09-242-202A-28	Sequence 28, Appl
3	1583.2	87.6	2308	16	US-09-242-202-29	Sequence 29, Appl
4	1583.2	87.6	2308	16	US-09-242-202A-29	Sequence 29, Appl
5	1310.4	72.5	1547	16	US-09-242-202-27	Sequence 27, Appl
6	1310.4	72.5	1547	16	US-09-242-202A-27	Sequence 27, Appl
7	607.4	33.6	7252	16	US-09-238-356-27	Sequence 27, Appl
8	604.2	33.4	7073	8	US-08-480-120-15	Sequence 15, Appl
9	604.2	33.4	7864	8	US-08-480-120-20	Sequence 20, Appl
10	603	33.4	608	16	US-09-242-202-26	Sequence 26, Appl
11	603	33.4	608	16	US-09-242-202A-26	Sequence 26, Appl
12	603	33.4	2077	1	PCT-US99-06742-7	Sequence 7, Appl
13	603	33.4	2462	18	US-09-496-445-5	Sequence 5, Appl
14	603	33.4	2462	44	US-60-118-478-5	Sequence 5, Appl
15	603	33.4	2577	3	US-07-919-535C-26	Sequence 26, Appl
16	603	33.4	2704	24	US-09-623-551-27	Sequence 27, Appl
17	603	33.4	2754	5	US-08-114-691-7	Sequence 7, Appl
18	603	33.4	2927	9	US-08-533-717-1	Sequence 1, Appl
19	603	33.4	2927	9	US-08-533-717A-1	Sequence 1, Appl
20	603	33.4	2927	15	US-09-142-481-14	Sequence 14, Appl
21	603	33.4	2939	1	PCT-US94-10146-2	Sequence 2, Appl
22	603	33.4	2939	8	US-08-488-015-2	Sequence 2, Appl
23	603	33.4	2939	15	US-09-114-835-2	Sequence 2, Appl
24	603	33.4	3175	1	PCT-US99-06704-2	Sequence 2, Appl
25	603	33.4	3175	25	US-09-646-075-2	Sequence 2, Appl
26	603	33.4	3301	8	US-08-447-430-42	Sequence 42, Appl
27	603	33.4	3301	17	US-09-342-673-42	Sequence 42, Appl
28	603	33.4	3327	21	US-09-546-411A-56	Sequence 56, Appl
29	603	33.4	3327	21	US-09-547-939A-56	Sequence 56, Appl
30	603	33.4	3327	45	US-09-628-730-56	Sequence 56, Appl
31	603	33.4	3327	45	US-60-128-830-56	Sequence 56, Appl
32	603	33.4	3331	15	US-09-148-920-1	Sequence 1, Appl
33	603	33.4	3331	15	US-09-148-920-1	Sequence 1, Appl
34	603	33.4	3344	18	US-09-422-569-8	Sequence 8, Appl
35	603	33.4	3369	21	US-09-546-411A-48	Sequence 48, Appl
36	603	33.4	3369	21	US-09-547-939A-48	Sequence 48, Appl
37	603	33.4	3369	45	US-09-628-730-48	Sequence 48, Appl
38	603	33.4	3369	45	US-60-128-830-48	Sequence 48, Appl
39	603	33.4	3420	15	US-09-643-192-12	Sequence 12, Appl
40	603	33.4	3421	15	US-09-130-115-31	Sequence 31, Appl
41	603	33.4	3423	8	US-08-447-430-40	Sequence 40, Appl
42	603	33.4	3423	17	US-09-342-673-40	Sequence 40, Appl
43	603	33.4	3423	17	US-09-342-673-40	Sequence 40, Appl

C	42	603	33.4	3423	18	US-09-485-737A-86	Sequence 86, Appl
C	43	603	33.4	3423	18	US-09-485-737B-86	Sequence 86, Appl
		603	33.4	3440	3	US-07-671-058-7	Sequence 7, Appl
		603	33.4	3462	29	US-09-742-373-10	Sequence 10, Appl

ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	100.0%;	Score 1807;	DB 16;	Length 1807;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1807;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGTACTCGCCACCATGGCGCGGATTCCTTTATCACTGATGAAGTTGGTGACATATTATGTT	60		
Db	1	GGTACTCGCCACCATGGCGCGGATTCCTTTATCACTGATGAAGTTGGTGACATATTATGTT	60		
QY	61	TATCACTGATAAAGTGTCAAGCATGCACAAAGTTGCAGCCGAATACAGTATCCGTGCCGG	120		
Db	61	TATCACTGATAAAGTGTCAAGCATGCACAAAGTTGCAGCCGAATACAGTATCCGTGCCGG	120		
QY	121	CCCTGGACTGTTGAACGAGGTGCGGCTAGACGGTCTCAGCACACGCAAACTGCCGAAACG	180		
Db	121	CCCTGGACTGTTGAACGAGGTGCGGCTAGACGGTCTCAGCACACGCAAACTGCCGAAACG	180		

181	QY	GTGTGGGGGTGCAGCAGCGCGCGCTTTACTGGCACATTCAGGAACAAGCGCGCCTTAAGG	240
181	Db		
181	QY	GTGTGGGGGTGCAGCAGCGCGCGCTTTACTGGCACATTCAGGAACAAGCGCGCGCCTTAAGG	240
181	Db		
241	QY	GCCATATGCTCAGTGGATGCTTGACCCACAGCGGGGATGAGGGGAGACCTGTAGTCAGAG	300
241	Db		
241	QY	GCCATATGCTCAGTGGATGCTTGACCCACAGCGGGGATGAGGGGAGACCTGTAGTCAGAG	300
241	Db		
301	QY	CCCCGGGCAGCACAGGCCAATGCCCGCTCTCCCTGCAAGATGATGATGAGTGCCTC	360
301	Db		
301	QY	CCCCGGGCAGCACAGGCCAATGCCCGCTCTCCCTGCAAGATGATGATGAGTGCCTC	360
301	Db		
361	QY	TCCTGGCCCTGGAAGTTCGCACTCCAGTGCACCACAGCCCTGTCTTAATAAAATTAAGTT	420
361	Db		
361	QY	TCCTGGCCCTGGAAGTTCGCACTCCAGTGCACCACAGCCCTGTCTTAATAAAATTAAGTT	420
361	Db		
421	QY	GCATCATTTTGTCTGACTAGTGTCTCTATAATATATTAAGCTTGATATCGAATTCCTTT	480
421	Db		
421	QY	GCATCATTTTGTCTGACTAGTGTCTCTATAATATATTAAGCTTGATATCGAATTCCTTT	480
421	Db		
481	QY	CGGACTTTTGAAGTATGTTGGTGGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC	540
481	Db		
481	QY	CGGACTTTTGAAGTATGTTGGTGGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC	540
481	Db		
541	QY	AGATTTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACCACGGGTAAATGCTTTTACT	600
541	Db		
541	QY	AGATTTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACCACGGGTAAATGCTTTTACT	600
541	Db		
601	QY	GGCCTGCTCCCTTATCGGGAAAGCGGGCGCATATATCAATGACGCCGCGCTGTAAGT	660
601	Db		
601	QY	GGCCTGCTCCCTTATCGGGAAAGCGGGCGCATATATCAATGACGCCGCGCTGTAAGT	660
601	Db		
661	QY	GTTAGCTTGAGAAAGAAATTCTCTGACGCCGCCGCTGTCTGGCGTTTTTCCATAGGCTCC	720
661	Db		
661	QY	GTTAGCTTGAGAAAGAAATTCTCTGACGCCGCCGCTGTCTGGCGTTTTTCCATAGGCTCC	720
661	Db		
721	QY	GCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCGACAG	780
721	Db		
721	QY	GCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCGACAG	780
721	Db		
781	QY	GACTATAAGATTACGACGCTTTCCCTCGTGAAGCTCCCTCGTGCCTCTCCTGTTCCCGA	840
781	Db		
781	QY	GACTATAAGATTACGACGCTTTCCCTCGTGAAGCTCCCTCGTGCCTCTCCTGTTCCCGA	840
781	Db		
841	QY	CCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGGCTGGCGCTTCTC	900
841	Db		
841	QY	CCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGGCTGGCGCTTCTC	900
841	Db		
901	QY	AATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTTCGCTCAAGCTGGGCTGTG	960
901	Db		
901	QY	AATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTTCGCTCAAGCTGGGCTGTG	960
901	Db		
961	QY	TGCACGAACCCCGCTTACGCCGACCGCTCGCCCTTATCCGGTAACTATCGTCTTGAGT	1020
961	Db		
961	QY	TGCACGAACCCCGCTTACGCCGACCGCTCGCCCTTATCCGGTAACTATCGTCTTGAGT	1020
961	Db		
1021	QY	CCAAACCGGTGAGACACAGACTTATGCCACTGGCAGCAGCCACTGTAAAGAGATTAGCA	1080
1021	Db		
1021	QY	CCAAACCGGTGAGACACAGACTTATGCCACTGGCAGCAGCCACTGTAAAGAGATTAGCA	1080
1021	Db		
1081	QY	GAGCAGGATGTGAGCGGTGTACAGAGTCTCTTCAAGTGGTGGCCTAACTACGGCTACA	1140
1081	Db		
1081	QY	GAGCAGGATGTGAGCGGTGTACAGAGTCTCTTCAAGTGGTGGCCTAACTACGGCTACA	1140
1081	Db		
1141	QY	CTAGAAGGACAGTATTTGGTATCTGGCTCTGCTGAAGCAGTTACCTTCGGAANAAGAG	1200
1141	Db		
1141	QY	CTAGAAGGACAGTATTTGGTATCTGGCTCTGCTGAAGCAGTTACCTTCGGAANAAGAG	1200
1141	Db		
1201	QY	TTGGTAGCTCTTCATCCGCAACAACACCCGCTGTAGCGGTGGTTTTTTTGTGTTGCA	1260
1201	Db		
1201	QY	TTGGTAGCTCTTCATCCGCAACAACACCCGCTGTAGCGGTGGTTTTTTTGTGTTGCA	1260
1201	Db		
1261	QY	AGCAGCAGATTACGCGCAGAAAAAAGGATCTCGGGGATTCGGGAGCTCCGACAGCGGTT	1320

Db 781 GACTATAAGATACACAGCGTTTCCCTTGGAGCTCCCTCGTGCCTCTCTCTGTTCCGA 840
QY 841 CCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCTGGCGTTTCTC 900
Db 841 CCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCTGGCGTTTCTC 900
QY 901 AATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTCGTCCAAAGCTGGCGTGTG 960
Db 901 AATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTCGTCCAAAGCTGGCGTGTG 960
QY 961 TGCACGAACCCCGCTTACGCGCGACCGCTGGCGCTTATCCGCTAACTATCGTCTGTAGT 1020
Db 961 TGCACGAACCCCGCTTACGCGCGACCGCTGGCGCTTATCCGCTAACTATCGTCTGTAGT 1020
QY 1021 CCAACCCGGTAAAGACAGCTATTCGCCACTGGCAGCAGCCACTGGTAAACAGATTAGCA 1080
Db 1021 CCAACCCGGTAAAGACAGCTATTCGCCACTGGCAGCAGCCACTGGTAAACAGATTAGCA 1080
QY 1081 GACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGGCTACA 1140
Db 1081 GACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGGCTACA 1140
QY 1141 CTAGAAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAG 1200
Db 1141 CTAGAAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAG 1200
QY 1201 TTGGTAGCTCTTATCGCGGCAACCAACACCCCTGGTAGCGGTGCTTTTGTGTTGCA 1260
Db 1201 TTGGTAGCTCTTATCGCGGCAACCAACACCCCTGGTAGCGGTGCTTTTGTGTTGCA 1260
QY 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGATCCGGAGAGCTCCCAACCGCTT 1320
Db 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGATCCGGAGAGCTCCCAACCGCTT 1320
QY 1321 GGATGCATGATAGGGAAGAGGATTAAGTCTGTAATGAATAGCAGGAACTTTGAGA 1380
Db 1321 GGATGCATGATAGGGAAGAGGATTAAGTCTGTAATGAATAGCAGGAACTTTGAGA 1380
QY 1381 CTCAGTCACTCAGTCACTAATAAGACTCAGTCACTCTGATCTGCTGCTTAACTGGCACT 1440
Db 1381 CTCAGTCACTCAGTCACTAATAAGACTCAGTCACTCTGATCTGCTGCTTAACTGGCACT 1440
QY 1441 CCTTGTGTTCCCAAGAACGGCTTCCTGCTCTGAGGAGGACCCCTTCCCTGGGAAGT 1500
Db 1441 CCTTGTGTTCCCAAGAACGGCTTCCTGCTCTGAGGAGGACCCCTTCCCTGGGAAGT 1500
QY 1501 AAACTAAGATGTCACGAGAGAAATTTTCCACCATTTGCTTGTGCTTAAAGAGGAAC 1560
Db 1501 AAACTAAGATGTCACGAGAGAAATTTTCCACCATTTGCTTGTGCTTAAAGAGGAAC 1560
QY 1561 TGATGAGTCACTCTAGATGAGAGCAGTCAAGGAGAGACAGAGACTCGAATTTCCGGA 1620
Db 1561 TGATGAGTCACTCTAGATGAGAGCAGTCAAGGAGAGACAGAGACTCGAATTTCCGGA 1620
QY 1621 GCTATTTTCAGTCTTCTTTTCCGTTTGTGCAATTTCACTTATGATACCGGCAATGCTG 1680
Db 1621 GCTATTTTCAGTCTTCTTTTCCGTTTGTGCAATTTCACTTATGATACCGGCAATGCTG 1680
QY 1681 GTTGCTATTTTGAACCTCCCTTAGGGGATGCCCCCTCACTGCCCCCTATAAAGGCCAG 1740
Db 1681 GTTGCTATTTTGAACCTCCCTTAGGGGATGCCCCCTCACTGCCCCCTATAAAGGCCAG 1740
QY 1741 CCTGAGCTGCAGAGATTCCTGCAGAGGATCAAGACAGCAGCTGGAGCTCAGAGGCTC 1800
Db 1741 CCTGAGCTGCAGAGATTCCTGCAGAGGATCAAGACAGCAGCTGGAGCTCAGAGGCTC 1800
QY 1801 TCCACACA 1807
Db 1801 TCCACACA 1807

; Sequence 29, Application US/09242202
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242.202
; FILING DATE: 01-Nov-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2308 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-242-202-29

Query Match 87.6%; Score 1583.2; DB 16; Length 2308;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1585; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 220 GAACAAGCGGGCCCTTAAGGGCCATATGGTGAAGTGGATGCTTGACCCAGCGGGGAT 279
Db 721 GTACAAGTGAAGCCCTTAAGGGCCATATGGTGAAGTGGATGCTTGACCCAGCGGGGAT 780
QY 280 GGGGAGACCTGTAGTCAGAGCCCCCGGCGAGCAGACAGGCAATGCCCGCTCTTCCCTGC 339
Db 781 GGGGAGACCTGTAGTCAGAGCCCCCGGCGAGCAGACAGGCAATGCCCGCTCTTCCCTGC 840
QY 340 AGGATGAGTAGTGAAGTGGCTCTCTCTGGCCCTTGAAGTGGCACTCCAGTCCAGCCAGCC 399
Db 841 AGGATGAGTAGTGAAGTGGCTCTCTCTGGCCCTTGAAGTGGCACTCCAGTCCAGCCAGCC 900
QY 400 TTGTCTTAATAAATAAGTGGTATCATTTTGTCTGACTAGGTGCTCTCTATAATATAT 459
Db 901 TTGTCTTAATAAATAAGTGGTATCATTTTGTCTGACTAGGTGCTCTCTATAATATAT 960
QY 460 AAGCTTGATATGAATTTCTTTTCGAGCTTTTGAAGTGAATGATGGTGGGGAAGATTGCA 519
Db 961 AAGCTTGATATGAATTTCTTTTCGAGCTTTTGAAGTGAATGATGGTGGGGAAGATTGCA 1020
QY 520 ACCTTCGAAGTCGATGACGCGCAGATTTAGAGTCTGTCTCTTTGGCGCTCGGGAACCCC 579
Db 1021 ACCTTCGAAGTCGATGACGCGCAGATTTAGAGTCTGTCTCTTTGGCGCTCGGGAACCCC 1080

Fri Jan 18 08:27:47 2002

QY 340 AGGATGAGTGTAGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACAGCC 399
DB 841 AGGATGAGTGTAGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACAGCC 900
QY 400 TTGCTCTAATAAATAAAGTTAGTTCATTTTGTCTGACTAGGTCTCTCTATAATATTAT 459
DB 901 TTGCTCTAATAAATAAAGTTAGTTCATTTTGTCTGACTAGGTCTCTCTATAATATTAT 960
QY 460 AAGCTTGATATCGAATTTCTTGGACATTTTGAAGTGTAGTGTGGGGGAGGATTCGA 519
DB 961 AAGCTTGATATCGAATTTCTTGGACATTTTGAAGTGTAGTGTGGGGGAGGATTCGA 1020
QY 520 ACCCTCGAAGTCGATGAGCGAGATTTAGAGTCTGCTCCCTTGGCCGCTCGGGAACCCC 579
DB 1021 ACCCTCGAAGTCGATGAGCGAGATTTAGAGTCTGCTCCCTTGGCCGCTCGGGAACCCC 1080
QY 580 ACCAGGGTAAATGCTTTTACTGGCTGCTCCCTTATCGGGAGCGGGCGCATCATATCA 639
DB 1081 ACCAGGGTAAATGCTTTTACTGGCTGCTCCCTTATCGGGAGCGGGCGCATCATATCA 1140
QY 640 AATGACGCGCCGCTGTAAAGTGTAGTGTGAGAAAGATTTCTGACGCGCCGCGGTTCG 699
DB 1141 AATGACGCGCCGCTGTAAAGTGTAGTGTGAGAAAGATTTCTGACGCGCCGCGGTTCG 1200
QY 700 TGGCGCTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGAGCTCAAGTC 759
DB 1201 TGGCGCTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGAGCTCAAGTC 1260
QY 760 AGAGTGGCGAAACCGACGAGCATATAAGATACAGCGCTTCCCGCTGGAAGCTCCC 819
DB 1261 AGAGTGGCGAAACCGACGAGCATATAAGATACAGCGCTTCCCGCTGGAAGCTCCC 1320
QY 820 TGTGCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGCGGCTTCTCCCTT 879
DB 1321 TGTGCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGCGGCTTCTCCCTT 1380
QY 880 CGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCCGTTGAGTGC 939
DB 1381 CGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCCGTTGAGTGC 1440
QY 940 TTCGCTCCAAGCTGGCGTGTGTCAGCAACCCCGCTTACGCCCGACCGCTGGCCTTAT 999
DB 1441 TTCGCTCCAAGCTGGCGTGTGTCAGCAACCCCGCTTACGCCCGACCGCTGGCCTTAT 1500
QY 1000 CCGGTAACTATCTCTGAGTCCAAACCGGTAGACACACACTTATCCCACTGGCAGCAG 1059
DB 1501 CCGGTAACTATCTCTGAGTCCAAACCGGTAGACACACACTTATCCCACTGGCAGCAG 1560
QY 1060 CCACTGTAACAGGATAGCAGACGAGGTATGTCGGGTGCTACAGAGTCTTGAAGT 1119
DB 1561 CCACTGTAACAGGATAGCAGACGAGGTATGTCGGGTGCTACAGAGTCTTGAAGT 1620
QY 1120 GGTGGCCCTAAGCTGCTACACTAGAGGACACTATTTGGTATCTCGGCTCTCTGAAGC 1179
DB 1621 GGTGGCCCTAAGCTGCTACACTAGAGGACACTATTTGGTATCTCGGCTCTCTGAAGC 1680
QY 1180 CAGTTACTCTCGGAAAAGAGTTGGTACCTTGTATCCGGCAACCAACCCGCTGGTA 1239
DB 1581 CAGTTACTCTCGGAAAAGAGTTGGTACCTTGTATCCGGCAACCAACCCGCTGGTA 1740
QY 1240 GCGGTGTTTTTTTGGTGAAGCAGAGATTTAGCGCCAGAAAAGAGATCTGGGGAT 1299
DB 1741 GCGGTGTTTTTTTGGTGAAGCAGAGATTTAGCGCCAGAAAAGAGATCTGGGGAT 1800
QY 1300 CCGGAGAGCTCCCAACCGTTGGATGATGATGAGGAAAGGAGTAAAGTCTGTAATG 1359
DB 1801 CCGGAGAGCTCCCAACCGTTGGATGATGATGAGGAAAGGAGTAAAGTCTGTAATG 1860
QY 1360 AATAAGCAGGAACCTTGAAGACTCAGTCACTAGTGTAGTAAATAAGACTCAGTACTTCT 1419
DB 1861 AATAAGCAGGAACCTTGAAGACTCAGTCACTAGTGTAGTAAATAAGACTCAGTACTTCT 1920
QY 1420 GATCCTGTCTTAAGTGCACCTCTTGTGTGTCACCAAGAAAGCGGCTTCTCTCTCAGG 1479

DB 1921 GATCCTGTCTTAAGTGCACCTCTTGTGTGTCACCAAGAAAGCGGCTTCTCTCTCAGG 1980
QY 1480 AGGACCCCTTCCCTGGAAGCTAAAGCTAAAGATGTGAGAGAGAAATTTTCCACCATGG 1539
DB 1981 AGGACCCCTTCCCTGGAAGCTAAAGCTAAAGATGTGAGAGAGAAATTTTCCACCATGG 2040
QY 1540 GTGCTTGGTCAAAAGAGAAACTGATGAGCTCACTCTAGATGAGAGAGCAGTGAAGGAG 1599
DB 2041 GTGCTTGGTCAAAAGAGAAACTGATGAGCTCACTCTAGATGAGAGAGCAGTGAAGGAG 2100
QY 1600 ACAGAGACTCGAATTTCCGAGCTATTTCAAGTATTTCTTTCCGTTTTTGTCAATTTCACT 1659
DB 2101 ACAGAGACTCGAATTTCCGAGCTATTTCAAGTATTTCTTTCCGTTTTTGTCAATTTCACT 2160
QY 1660 TATGATACCGGCCAATCTTGGTTGCTTATTTTGGAACTCCCTTAGGGGATGCCCTCA 1719
DB 2161 TATGATACCGGCCAATCTTGGTTGCTTATTTTGGAACTCCCTTAGGGGATGCCCTCA 2220
QY 1720 ACTGGCCCTATAAAGGCCAGCCTGAGCTGAGAGGATTTCTGAGAGGATCAAGACAGC 1779
DB 2221 ACTGGCCCTATAAAGGCCAGCCTGAGCTGAGAGGATTTCTGAGAGGATCAAGACAGC 2280
QY 1780 ACGTGGACCTCGACAGCCTCTCCACCA 1807
DB 2281 ACGTGGACCTCGACAGCCTCTCCACCA 2308

RESULT 5
US-09-242-202-27
; Sequence 27, Application US/09242202
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202
; FILING DATE: 01-NOV-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1547 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLFCULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-242-202-27

Query Match 72.5%; Score 1310.4; DB 16; Length 1547;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACCATGGCGGGATCTTTATCACTGATTAAGTTGGTGACATATATGTT 60
DB 1 GGTACCTGCCACCATGGCGGGATCTTTATCACTGATTAAGTTGGTGACATATATGTT 60
QY 61 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCGAGCCGAATACAGTGCCTGCGG 120
DB 61 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCGAGCCGAATACAGTGCCTGCGG 120
QY 121 CCCTGGACTGTTGAACGAGGTGCGGCTAGACGTCTGACGACGCAAACTGGCGAAG 180
DB 121 CCCTGGACTGTTGAACGAGGTGCGGCTAGACGTCTGACGACGCAAACTGGCGAAG 180
QY 181 GTTGGGGTGCACGAGCGCGCTTTACTGGCACTTCAGGAACAGCGGCGCTTAAGG 240
DB 181 GTTGGGGTGCACGAGCGCGCTTTACTGGCACTTCAGGAACAGCGGCGCTTAAGG 240
QY 241 GCCATATGTTGAGTGGATGCTTTGACCCAGCGGGATGGGGAGACCTGTAGTCAGAG 300
DB 241 GCCATATGTTGAGTGGATGCTTTGACCCAGCGGGATGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCCGGGACGACAGGCCAATCCCGTCTTCCCTGCGAGGATGAGTAGTGCGCTC 360
DB 301 CCCCCGGGACGACAGGCCAATCCCGTCTTCCCTGCGAGGATGAGTAGTGCGCTC 360
QY 361 TCCTGGCCCTGGAAGTTGCCACTCCAGTGCACCGCCCTTGTCTTAATAAATTAAGTT 420
DB 361 TCCTGGCCCTGGAAGTTGCCACTCCAGTGCACCGCCCTTGTCTTAATAAATTAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGTGTCTCTATAATATTAAGTGTGATATCGAATTTCT 480
DB 421 GCATCATTTTGTCTGACTAGTGTCTCTATAATATTAAGTGTGATATCGAATTTCT 480
QY 481 CGCAGTTTGAAGATGATGTGTGGGGAGGATTCGAACCTTCGAGTCGATGCGGC 540
DB 481 CGCAGTTTGAAGATGATGTGTGGGGAGGATTCGAACCTTCGAGTCGATGCGGC 540
QY 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
DB 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
QY 601 GGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
DB 601 GGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
QY 661 GTTACGTTGAGAAGATTCCTGACGCGCGCGGTGCTGGGTTTTCCATAGGCTCC 720
DB 661 GTTACGTTGAGAAGATTCCTGACGCGCGCGGTGCTGGGTTTTCCATAGGCTCC 720
QY 721 GCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAG 780
DB 721 GCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAG 780
QY 781 GACTATAAAGATACAGGCGCTTTCCCGCTGGAAGCTCCCTGCGCTCCCTGTTCCGA 840
DB 781 GACTATAAAGATACAGGCGCTTTCCCGCTGGAAGCTCCCTGCGCTCCCTGTTCCGA 840
QY 841 CCCTGGCGCTTACGGATACCTGCGGCTTTCTCCCTGCGGAAAGGCTGGCGTTTCTC 900
DB 841 CCCTGGCGCTTACGGATACCTGCGGCTTTCTCCCTGCGGAAAGGCTGGCGTTTCTC 900
QY 901 AATGCTCACGCTGTAGTATCTCAGTTCGCTGAGTGTAGTGTGCTCAAGCTGGGCTGTG 960
DB 901 AATGCTCACGCTGTAGTATCTCAGTTCGCTGAGTGTAGTGTGCTCAAGCTGGGCTGTG 960
QY 961 TGCACGAACCCCGTTACGCGGACCGCTGCGCTTATCCGCTAAGTATCGCTTTCGAGT 1020
DB 961 TGCACGAACCCCGTTACGCGGACCGCTGCGCTTATCCGCTAAGTATCGCTTTCGAGT 1020

DB 961 TGCACGAACCCCGTTACGCGGACCGCTTATCCGCTAAGTATCGCTTTCGAGT 1020
QY 1021 CCAACCCGGTAAACACAGGACTTATCGCCACTGCGCAGCCACTGGTAAACAGGATTAGCA 1080
DB 1021 CCAACCCGGTAAACACAGGACTTATCGCCACTGCGCAGCCACTGGTAAACAGGATTAGCA 1080
QY 1081 GAGCGAGTATGTAGCGGCTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACA 1140
DB 1081 GAGCGAGTATGTAGCGGCTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACA 1140
QY 1141 CTAGAAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGGAG 1200
DB 1141 CTAGAAGCAGATATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGGAG 1200
QY 1201 TTGGTAGCTTTCATCCGCGCAAAACAAACACCGCTGCTGAGCGGTGTTTGTGTTGCA 1260
DB 1201 TTGGTAGCTTTCATCCGCGCAAAACAAACACCGCTGCTGAGCGGTGTTTGTGTTGCA 1260
QY 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGATCCGGAGAGCTCCC 1312
DB 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGATCCGGAGAGCTCAC 1312

RESULT 6

US-09-242-202a-27
; Sequence 27, Application US/09242202a

; GENERAL INFORMATION:

; APPLICANT: Nelson, Edward L.

; TITLE OF INVENTION: NOVEL VECTOR FOR

; POLYNUCLEOTIDE VACCINES

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: MS WORD 97

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/242,202a

; FILING DATE: 20-Apr-2000

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US60/023931

; FILING DATE: 14-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: KATHRYN M. BROWN

; REGISTRATION NUMBER: 34556

; REFERENCE/DOCKET NUMBER: 2026-4236US1

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1547 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: No

; ANTI-SENSE: No

; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-242-202a-27

Query Match 72.5%; Score 1310.4; DB 16; Length 1547;
Best Local Similarity 99.9%; Pred. No. 0;

QY 1020 TCCAAACCCGGTAAACACAGCACTTATCCCACTGGCAGCAGCCACTGTAAACAGGATTAGC 1079
Db 5422 tccaaacccggtaagacagcaacttatcccaactggcagcagccactggttaacaggattagc 5481
QY 1080 AGAGCGAGGTATAGCGGTGTACAGAGTCTTGAAGTGGGCTAACTACGGCTAC 1139
Db 5482 agagcagggtatgagcggtgtacagagttcttgaagtggcgctaaactacggctac 5541
QY 1140 ACTAGAGGACACTATTGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGA 1199
Db 5542 accagaagagcagatttggatctgctgctctgctgaagccagttaccttcggaaaga 5601
QY 1200 GTTGGTAGTCTTGTATCCGGCAACAAACACCGCTGGTGAAGCGTGTGTTTGTGTTGC 1259
Db 5602 gttggtagctctgtatccggcaacaaacacccgcgtggtgaagccagttaccttcggaaaga 5661
QY 1260 AAGCAGAGATTACGGCGCAGAAAAAAGGATCT 1292
Db 5662 aagcagagattacgcgcagaaaaaaggatct 5694

RESULT 8
US-08-480-120-15
; Sequence 15, Application US/08480120
; GENERAL INFORMATION:
; APPLICANT: Joliffe, Linda K.
; APPLICANT: Zivin, Robert A.
; APPLICANT: Pulito, Virginia L.
; TITLE OF INVENTION: CDR-GRAFTED ANTI-TISSUE FACTOR
; TITLE OF INVENTION: ANTIBODIES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,120
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9598
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..717
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1111..1146
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1268..1594
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 1692..2012
US-08-480-120-15

Query Match 33.4%; Score 604.2; DB 8; Length 7073;
Best Local Similarity 97.9%; Pred. No. 1.8e-167;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 568 TGAGAAAGAAATTCCTGCAGCGCGCGGCTTGCCTGGCGCTTTTCCATAGGCTCCGCCCC 727
Db 2856 TAAGGAGAGAGCGCTGCACCTCGGCGCGGCTTGCCTGGCGCTTTTCCATAGGCTCCGCCCC 2915
QY 728 TGACGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATA 787
Db 2916 TGACGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATA 2975
QY 788 AAGATACCGAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTCCGACCCGTC 847
Db 2976 AAGATACCGAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTCCGACCCGTC 3035
QY 848 GCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
Db 3036 GCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTC 3095
QY 908 ACGCTGTAGGTATCTCAAGTTCGGGTGTAGGTCTCGCTCCAAAGCTGGGCTGTGTGCACGA 967
Db 3096 ACGCTGTAGGTATCTCAAGTTCGGGTGTAGGTCTCGCTCCAAAGCTGGGCTGTGTGCACGA 3155
QY 968 ACCCCCGCTTCAGCCCGGAGCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACC 1027
Db 3156 ACCCCCGCTTCAGCCCGGAGCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACC 3215
QY 1028 GGTAAAGACAGCACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGCGAG 1087
Db 3216 GGTAAAGACAGCACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGCGAG 3275
QY 1088 GTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG 1147
Db 3276 GTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG 3335
QY 1148 GACAGTATTTGGTATCTGCGCTCTGTGTAAGCCAGTTACCTTCGGAAGAGAGTTGGTAG 1207
Db 3336 GACAGTATTTGGTATCTGCGCTCTGTGTAAGCCAGTTACCTTCGGAAGAGAGTTGGTAG 3395
QY 1208 CTCTTGATCCGGCAAAACAAACCACCGCTGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 1267
Db 3396 CTCTTGATCCGGCAAAACAAACCACCGCTGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 3455
QY 1268 GATTACGCGCAGAAAAAAGGATCT 1292
Db 3456 GATTACGCGCAGAAAAAAGGATCT 3480

RESULT 9
US-08-480-120-20
; Sequence 20, Application US/08480120
; GENERAL INFORMATION:
; APPLICANT: Joliffe, Linda K.
; APPLICANT: Zivin, Robert A.
; APPLICANT: Pulito, Virginia L.
; TITLE OF INVENTION: CDR-GRAFTED ANTI-TISSUE FACTOR
; TITLE OF INVENTION: ANTIBODIES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Fri Jan 18 08:27:47 2002

us-09-242-202a-28.rnbp

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,120
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9598
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 7864 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: CDS
LOCATION: 9..711
US-08-480-120-20

Query Match 33.4%; Score 604.2; DB 8; Length 7864;
Best Local Similarity 97.9%; Pred. No. 1.8e-167;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAGAAATTCCTGAGCGCGCGCTGCTGGCGTTTTCATAGGCTCCGCCCGC 727
DB 1255 TAAGGAGAGCGTCACTCGGCGCGCTGCTGGCGTTTTCATAGGCTCCGCCCGC 1314
QY 728 TGAGAGCATCACAAATGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 787
DB 1315 TGAGAGCATCACAAATGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 1374
QY 788 AAGTACAGCGGTTTCCCTCGGAAGTCCCTCGGAGCGTGGCGTCTTCTCAATGCTC 907
DB 1375 AAGTACAGCGGTTTCCCTCGGAAGTCCCTCGGAGCGTGGCGTCTTCTCAATGCTC 1434
QY 848 GCTTACCGGATACCTGCTCGGCTTCTCCCTTCGGGAGCGTGGCGTCTTCTCAATGCTC 907
DB 1435 GCTTACCGGATACCTGCTCGGCTTCTCCCTTCGGGAGCGTGGCGTCTTCTCAATGCTC 1494
QY 908 AGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967
DB 1495 AGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
QY 968 ACCCGCGTTACCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027
DB 1555 ACCCGCGTTACCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1614
QY 1028 GGTAAAGACAGCACTTATGCCACTGGCAGCAGCCACTGGTAAACAGGATACAGAGCGAG 1087
DB 1615 GGTAAAGACAGCACTTATGCCACTGGCAGCAGCCACTGGTAAACAGGATACAGAGCGAG 1674
QY 1088 GTATGAGCGGCTGTACAGAGTTCCTGAAGTGGTGGCTTACCTACGCTACCTAGAGAG 1147
DB 1675 GTATGAGCGGCTGTACAGAGTTCCTGAAGTGGTGGCTTACCTACGCTACCTAGAGAG 1734
QY 1148 GACAGTATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207
DB 1735 GACAGTATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1794
QY 1208 CTCTTGATCCGGCAACAAACACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1267
DB 1795 CTCTTGATCCGGCAACAAACACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1854
QY 1268 GATTACCGCGAGAAAAAGGATCT 1292

DB 1855 GATTACCGCGAGAAAAAGGATCT 1879

RESULT 10
US-09-242-202-26
Sequence 26, Application US/09242202
GENERAL INFORMATION:
APPLICANT: Nelson, Edward L.
Nelson, Peter J.
TITLE OF INVENTION: NOVEL VECTOR FOR POLYNUCLEOTIDE VACCINES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/242,202
FILING DATE: 01-Nov-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US60/023931
FILING DATE: 14-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34556
REFERENCE/DOCKET NUMBER: 2026-4236US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-242-202-26

Query Match 33.4%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 1.3e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCGAGGTTTCCCGCT 749
DB 6 GCCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCGAGGTTTCCCGCT 65
QY 750 GCCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCGAGGTTTCCCGCT 809
DB 66 GCCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCGAGGTTTCCCGCT 125
QY 810 GGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 869
DB 126 GGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 185
QY 870 TTTCTCCCTTCGGGAAGCGTGGCGCTTTCATGCTCAGCTGTAGGTATCTCAGTTTCG 929
DB 186 TTTCTCCCTTCGGGAAGCGTGGCGCTTTCATGCTCAGCTGTAGGTATCTCAGTTTCG 245
QY 930 GTGTAGTTCGTTCCCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGC 989

Db 246 GTGAGGCTGCTTCCAGAGTGGGCTGTGTGCAGAACCCCGTTTCAGCCGACCCG 305
QY 990 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 1049
Db 306 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 365
QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 1109
Db 366 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 425
QY 1110 TTCTTGAAGTGTGGCTTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 1169
Db 426 TTCTTGAAGTGTGGCTTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 485
QY 1170 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTAGTCTTGTATCCGCGCAACAAACC 1229
Db 486 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTAGTCTTGTATCCGCGCAACAAACC 545
QY 1230 ACCGCTGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGCAAGAAAAGGA 1289
Db 546 ACCGCTGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGCAAGAAAAGGA 605
QY 1290 TCT 1292
Db 606 TCT 608

RESULT 11

US-09-242-202a-26
; Sequence 26, Application US/09242202a
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; TITLE OF INVENTION: NOVEL VECTOR FOR POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242, 202a
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-42360S1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 608 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-242-202a-26

Query Match 33.4%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 1.3e-167; Indels 0; Gaps 0;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCGCGCTTCTCGCGCTTTTTCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 749
Db 6 GCGCGCTTCTCGCGCTTTTTCATAGGCTCCGCCCGCTGACGAGCATCACAAAATCGA 65
QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACGCGGTTCCTCCCT 809
Db 66 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACGCGGTTCCTCCCT 125
QY 810 GGAAGCTCCCTCGTGGGCTCTCTGTTCGACCGCTCCCGCTACCGGATACCTGTCCGCC 869
Db 126 GGAAGCTCCCTCGTGGGCTCTCTGTTCGACCGCTCCCGCTACCGGATACCTGTCCGCC 185
QY 870 TTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCG 929
Db 186 TTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCG 245
QY 930 GTGTAGGTCTCGCTCCAAAGCTGGGCTGTGTGCAGAAACCCCGCTTCAGCCCGACCCG 989
Db 246 GTGTAGGTCTCGCTCCAAAGCTGGGCTGTGTGCAGAAACCCCGCTTCAGCCCGACCCG 305
QY 990 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 1049
Db 306 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 365
QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 1109
Db 366 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 425
QY 1110 TTCTTGAAGTGTGGCTTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 1169
Db 426 TTCTTGAAGTGTGGCTTAACATATCTGTAGTCCAAACCCGGTAAGACACGACTATTCGCCA 485
QY 1170 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTAGTCTTGTATCCGCGCAACAAACC 1229
Db 486 CTGCTGAAGCCAGTACCTTCGGAAGAGAGTGTAGTCTTGTATCCGCGCAACAAACC 545
QY 1230 ACCGCTGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGCAAGAAAAGGA 1289
Db 546 ACCGCTGTAGCGGTGGTGTGTTTGTGCAAGCAGCAGATACCGCGCAAGAAAAGGA 605
QY 1290 TCT 1292
Db 606 TCT 608

RESULT 12

PCT-US99-06742-7
; Sequence 7, Application PC/TUS9906742
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/06742
; CURRENT FILING DATE: 1999-03-28
; EARLIER APPLICATION NUMBER: 60/079,792
; EARLIER FILING DATE: 1998-03-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Word97
; SEQ ID NO 7
; LENGTH: 2077
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: paslib

PCT-US99-06742-7

Query Match 33.4%; Score 603; DB 1; Length 2077;
 Best Local Similarity 100.0%; Pred. No. 2.3e-167;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCGCTTTTCCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGA 749
 Db 1437 gcgcggttgcgtttttccataggtccgccccctgacgagcatcaaaaaatcga 1496
 QY 750 CCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAGATACAGCGCTTCCGCCCT 809
 Db 1497 cgcctcaagtgcagagtggtgcgaacccgacagactataaagataccagcggttccccct 1556
 QY 810 GGAAGCTCCCTGCGCT 869
 Db 1557 ggaagctccctgcgtct 1616
 QY 870 TTTTCCCTTCGGGAAGCTGGCGCTTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTGC 929
 Db 1617 ttctcccttcgggaagcgtggcgctttctcaatgctcagcgtgtaggtatctcagttcg 1676
 QY 930 GTGTAGTCTGCTTCCAAAGCTGGGCTGTGTGACGACGCCGCCCTTCAGCGCGACCGC 989
 Db 1677 gtgtaggtctgctccaaagctggcgtgtgtgacgaaccccggttcagccgacgcg 1736
 QY 990 TGGCGCTTATCCGGTAACTATCGCTTTGAGTCCAAACCCGGTAAGACAGCACTATCGCCA 1049
 Db 1737 tggccttatccggttaactatcgtcttgatccaaacccggtaagacagacttaacgcca 1796
 QY 1050 CTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGAGCGGCTGTACAGAG 1109
 Db 1797 ctggcagcagccactggttaacaggaattagcagagcgaggtatgtaggcggtgtacagag 1856
 QY 1110 TTTCTGAAGTGTGGCTTAACCTACGCTACACTAGAGACAGATTTGGTATCTCGGCT 1169
 Db 1857 ttctgaagtgtggcctaactacgctacactagagagacagattttggtatctgcgt 1916
 QY 1170 CTGCTGAAGCCAGTACCTTCGAAAAGGTGGTGTAGCTTTGATCCGGCAAAACACC 1229
 Db 1917 ctgctgaagccagtaccttcgaaaaggtggtagctttgattccttgatccggcaaaacacc 1976
 QY 1230 ACCGCTGTAGCGGTGGT 1289
 Db 1977 accgctgtgagcggt 2036

RESULT 13
 US-09-496-445-5
 ; Sequence 5, Application US/09496445
 ; GENERAL INFORMATION:
 ; APPLICANT: McNeish, John D.
 ; APPLICANT: Ahljinian, Michael K
 ; TITLE OF INVENTION: Transgenic Animals Expressing Human p25
 ; FILE REFERENCE: PC10142A
 ; CURRENT APPLICATION NUMBER: US/09/496,445
 ; CURRENT FILING DATE: 2000-02-02
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 2462
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-496-445-5

Query Match 33.4%; Score 603; DB 18; Length 2462;
 Best Local Similarity 100.0%; Pred. No. 2.5e-167;

Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCGCTTTTCCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGA 749
 Db 359 gcgcggttgcgtttttccataggtccgccccctgacgagcatcaaaaaatcga 418
 QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAGATACAGCGGTTCGCCCT 809
 Db 419 cgcctcaagtgcagagtggtgcgaacccgacagactataaagataccagcggttccccct 478
 QY 810 GGAAGCTCCCTGCGCT 869
 Db 479 ggaagctccctgcgtct 538
 QY 870 TTTTCCCTTCGGGAAGCTGGGCTTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTGC 929
 Db 539 ttctcccttcgggaagcgtggcgctttctcaatgctcagcgtgtaggtatctcagttcg 598
 QY 930 GTGTAGTCTGCTTCCAAAGCTGGGCTGTGTGACGACGCCGCCCTTCAGCGCGACCGC 989
 Db 599 gtgtaggtctgctccaaagctggcgtgtgtgacgaaccccggttcagccgacgcg 658
 QY 990 TGGCGCTTATCCGGTAACTATCGCTTTGAGTCCAAACCCGGTAAGACAGCACTATCGCCA 1049
 Db 659 tggccttatccggttaactatcgtcttgatccaaacccggtaagacagacttaacgcca 718
 QY 1050 CTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGAGCGGCTGTACAGAG 1109
 Db 719 ctggcagcagccactggttaacaggaattagcagagcgaggtatgtaggcggtgtacacag 778
 QY 1110 TTTCTGAAGTGTGGCTTAACCTACGCTACACTAGAGACAGATTTGGTATCTCGGCT 1169
 Db 779 ttctgaagtgtggcctaactacgctacactagagagacagattttggtatctgcgt 838
 QY 1170 CTGCTGAAGCCAGTACCTTCGAAAAGGTGGTGTAGCTTTGATCCGGCAAAACACC 1229
 Db 839 ctgctgaagccagtaccttcgaaaaggtgtgtagctttgattccttgatccggcaaaacacc 898
 QY 1230 ACCGCTGTAGCGGTGGT 1289
 Db 899 accgctgtgagcggt 958
 QY 1290 TCT 1292
 Db 959 tct 961

Query Match 33.4%; Score 603; DB 44; Length 2462;
 Best Local Similarity 100.0%; Pred. No. 2.5e-167;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCGCTTTTCCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGA 749
 Db 359 gcgcggttgcgtttttccataggtccgccccctgacgagcatcaaaaaatcga 418


```

QY 750 CGCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCAAGCGGTTTCCCGCT 809
Db 419 cgcctcaagtcagagtgccgaaacccgacagactataaagataccagcggtttcccccct 478
QY 810 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCC 869
Db 479 ggaagctccctcgctggctctctctggtccgacctgctgcttaccgataccctgtccgcc 538
QY 870 TTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTCG 929
Db 539 ttctcccttcgggaaagcgctggctctctcaatgctcagcgctaggtatctcagttcg 598
QY 930 GTGTAGTCTGCTCGCTCAAGCTGGCGTGTGTCACGAAACCCCGCTTCAGCGGACCGC 989
Db 599 gtgtagtcgtcgtcccaagctgggctgtgtgacgaacccccgttaccgacccgcgc 658
QY 990 TGGCGCTTATCCCGTAACTACTGCTTGTGACGCAACCCGGTAAGACACGACTTATCGCCA 1049
Db 659 tgcgcttatccggttaactatcgtcttgagtcgaacccggttaagacacgacttatcgcca 718
QY 1050 CTGGCAGCAGCCACTGTAAACAGATTAGCAGGCGAGGTATGTAGCGGTGCTACAGAG 1109
Db 719 ctggcagcagccactgttaacagatttagcagagcgaggtatgtagcggtgctacagag 778
QY 1110 TTTCTGAAGTGGTGGCGCTAACTACGGCTACACTAGAGGACAGTATTGTTGTTATCTGCGCT 1169
Db 779 ttctgaagtgggtggcctaaactacgctacactagaagacagattgtgtatctgcgct 838
QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTGGTGTGATCTGATCCGCGCAACAAACC 1229
Db 839 ctgctgaagccagttaaccttcggaagaggttggtagctctgacgcgcaacaaacc 898
QY 1230 ACCGCTGTAGCGGTGTTTGTGTTGTTGCAACGACAGATTACGCGCAGAAAAAGGA 1289
Db 899 accgctgtagcggtggttgtgtgtgttcttctgcaagcagcagattacgcgcaaaaaagga 958
QY 1290 TCT 1292
Db 959 tct 961

```

```

RESULT 15
US-07-919-535C-26
; Sequence 26, Application US/07919535C
; GENERAL INFORMATION:
; APPLICANT: Haas, Werner
; APPLICANT: Hunziker, Willi
; TITLE OF INVENTION: SOLUBLE KIT LIGANDS
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: DOS 4.0
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/919,535C
; FILING DATE: 19920723
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91810609.7
; FILING DATE: 30-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P.
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: RAN 4105/142

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2577 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 121..126
; OTHER INFORMATION: /note= "Recognition site for
; OTHER INFORMATION: restriction enzyme NcoI"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 141..146
; OTHER INFORMATION: /note= "Recognition site for
; OTHER INFORMATION: restriction enzyme HindIII"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2577
; OTHER INFORMATION: /note= "Complete nucleotide
; OTHER INFORMATION: sequence of E. coli
; OTHER INFORMATION: expression plasmid
; OTHER INFORMATION: pOC 56/RBS II, NcoI"
; US-07-919-535C-26

Query Match 33.4%; Score 603; DB 3; Length 2577;
Best Local Similarity 100.0%; Pred. No. 2.5e-167; Indels 0; Gaps 0;
Matches 603; Conservative 0; Mismatches 0;

QY 690 GCCGCGTGTGCTGGCGTTTTTCATAGCTCGCGCCCGCTGACGAGCATCAGAAAAATCGA 749
Db 736 GCCGCGTGTGCTGGCGTTTTTCATAGCTCGCGCCCGCTGACGAGCATCAGAAAAATCGA 795
QY 750 CGCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCAAGCGGTTTCCCGCT 809
Db 796 CGCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACCAAGCGGTTTCCCGCT 855
QY 810 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCC 869
Db 856 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGCGGTTACCGGATACCTGTCCGCC 915
QY 870 TTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTCG 929
Db 916 TTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTCG 975
QY 930 GTGTAGTGTGCTCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGCGACCGC 989
Db 976 GTGTAGTGTGCTCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGCGACCGC 1035
QY 990 TGGCGCTTATCCGCTAACTATCGTCTGAGTCCAAACCCGTAAGACACGACTTATCGCCA 1049
Db 1036 TGGCGCTTATCCGCTAACTATCGTCTGAGTCCAAACCCGTAAGACACGACTTATCGCCA 1095
QY 1050 CTGGCAGCAGCCACTGTGTAACAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG 1109
Db 1096 CTGGCAGCAGCCACTGTGTAACAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG 1155
QY 1110 TTCTTGAAGTGGTGGCGCTTAACAGGCTACAGGCTACACTAGAAGGACAGTATTGTTGTTATCTGGCT 1169
Db 1156 TTCTTGAAGTGGTGGCGCTTAACAGGCTACAGGCTACACTAGAAGGACAGTATTGTTGTTATCTGGCT 1215
QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTGGTGTGATCTGATCCGCGCAACAAACC 1229
Db 1216 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTGGTGTGATCTGATCCGCGCAACAAACC 1275
QY 1230 ACCGCTGTAGCGGTGTTTGTGTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAGGA 1289

```

|||||
Db 1276 ACCGCTAGCGGTGTTTTTTTTCGTTGCAAGCAGCAGATTACCGCAGAAAAAGGA 1335
Qy 1290 TCT 1292
|||
Db 1336 TCT 1338

Search completed: January 17, 2002, 15:14:36
Job time: 22788 sec